Winter 2015

Educational Pathways of Teachers and the Effects on Students' Performance on High-Stakes Testing in Texas

Patrizia Antoinette Grigsby
Old Dominion University

EDUCATIONAL PATHWAYS OF TEACHERS AND THE EFFECTS ON STUDENTS' PERFORMANCE ON HIGH-STAKES TESTING IN TEXAS

by

Patrizia Antoinette Grigsby

B.S. June 1990, University of Tennessee

M.A. May 2004, University of Saint Mary's

Ed.S. December 2008, Old Dominion University

A Dissertation Submitted to the Faculty of Old Dominion University in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

EDUCATIONAL LEADERSHIP

OLD DOMINION UNIVERSITY

December 2015

Approved by:

William Owings (Chair)

John Ritz (Member)

Steve Myran (Member)

ABSTRACT

EDUCATIONAL PATHWAYS OF TEACHERS AND THE EFFECTS ON STUDENTS' PERFORMANCE ON HIGH-STAKES TESTING

Patrizia Antoinette Grigsby

Old Dominion University, 2015

Director: Dr. William Owings

Meeting the needs of placing "highly-qualified teachers" in the classroom is becoming more challenging every year. One option that many states are using involves certifying potential teaching candidates through alternative pathways, which differs from traditional certification programs. In this study, teacher certification routes were examined to determine if there is a difference in student performance on state standardized tests based on the teacher certification routes. This study compared student test scores of traditionally-certified teachers with test scores from students of alternatively-certified teachers on the State of Texas Assessments of Academic Readiness (STAAR) during the 2011-2012 administration of the test. The study findings revealed that alternatively-certified teachers' student scores were comparable to traditionally-certified teachers' student scores on the STAAR tests after three to five years of teaching experience and professional development. These findings can be a precursor to the efforts of other states evaluating whether traditionally-certified teachers and alternatively – certified teachers' student scores on state standardized tests are comparable when including other quantitative variables such as teacher performance on basic skill tests and

teacher quality data. This will ensure prospective teachers have the ability to provide the teaching quality needed to increase the probability of successful student achievement.

Dedication

This dissertation is dedicated to my family and friends. I thank them for their advice, support, and love throughout this process. I thank my mother for her prayers and constant reminders that anything worth achieving takes endurance, perseverance, and a positive attitude. I thank my husband and son for understanding my desire to pursue my Ph.D., being considerate, and showing me unconditional love.

Acknowledgements

I would like to acknowledge Dr. Owings for his support and guidance throughout my pursuit of my Ph.D. in Educational Leadership. He has constantly encouraged me to complete my coursework, be confident in obtaining my academic goals, and become an expert in the field of education. I am forever grateful for his encouraging words and unwavering support.

Page

Table of Contents

Page

LIST OF TABLES

TABLE OF FIGURES

Page

CHAPTER I

INTRODUCTION

Student achievement is used increasingly to measure teacher classroom effectiveness for evaluation purposes. At the same time, many states are exploring the process of licensing potential teacher candidates through alternative pathways. Alternate routes to teacher licensure provide opportunities for school divisions to hire potential teacher candidates who have met subject matter competency, but may not have studied education while in college (Feistritzer, 2008). In this study, teacher certification pathways and their effects on student performance on state standardized tests will be examined. When teachers obtain licensure, it indicates that they have earned formal approval for professional practice. These terms will be used interchangeably throughout the study.

History of Teacher Licensure

In the 19th century, teachers achieved licensure by taking a test administered by the local school division or county board of education. The number of education courses a teacher had attended did not matter when teachers pursued their teaching license. In 1898, only four states had established centralized state licensure (Kaplan & Owings, 2011). During the 19th century, the authority for licensing teachers passed from ecclesiastical to civil authorities (Angus, 2001; Feistritzer, 2008). At first, civil authorities were officials from the local government who issued teacher licensure based on the teacher candidates' moral character and performance on an oral or written test. At that time, the exam covered content knowledge that the teacher would teach. The

reasoning for this was to make sure that the teacher knew more than the older students (Feistritzer, 2008). In 1902, Superintendent Maxwell of the New York City School System was instrumental in insisting that teacher candidates from the Normal College would take the final city and state exams. Maxwell's actions were part of a larger plan that included standardized regulations for teacher candidates, which resulted in schools being able to certify courses taken by students (Fraser, 2007). The certification of coursework taken by the students gave professional organizations the ability to grant licensure to teachers.

In the late 19th and early 20th centuries, as the United States grew more industrialized, the responsibility for teacher professional training changed from the localities to the states. By 1933, 42 states had centralized state teacher licensure requirements. During this time, completion of teacher education programs and coursework became necessary to obtain licensure for prospective teachers (Kaplan & Owings, 2011).

Beginnings

The majority of states had a two- or three-tiered licensure process, with additional requirements associated with obtaining each type of license (Mitchell et al., 2001). In all, 31 states required an initial certification (which was valid for two or five years) and the attainment of a standard professional license based on the completion of additional coursework. Another 13 states offered an advanced certificate and three states offered a lifetime certificate at the advanced level. Most states required advanced degrees for additional licensure requirements. However, some states (i.e. California, North Carolina, and Connecticut) required demonstration of teaching competencies to acquire the next

level of licensure. Generally speaking, licensure requirements caused a reduction in the number of prospective teacher candidates. Hard-to-fill teaching positions in science, math, special education, and bilingual education existed in urban and rural areas. To address this concern, some school divisions offered a provisional licensure to individuals who had a college degree or a license from a different state.

Modern Era of Teaching Licensure

Each state currently has its own specific requirements for licensing teacher candidates. Most states require a combination of the following criteria: formal academic training, completion of an accredited teacher preparation program, completion of an internship or practicum, and passing the state's licensure exam. In 2002, the enactment of No Child Left Behind Act (NCLB) by the federal government specified new definitions for highly-qualified teachers. Highly qualified teachers were now required to possess a college degree and full state certification or licensure. Temporary, emergency, and provisional licensure did not fulfill the requirements to being considered highly qualified (Education Commission, 2002). Teachers demonstrated their content knowledge through the following proficiencies: elementary teachers passed a state test of numeracy and literacy; secondary teachers passed a rigorous test of their subject matter or area of concentration, and veteran teachers passed a state test or examination based on their subject matter knowledge (Education Commission, 2002).

Purpose of Licensure

Teacher licensure has long been regulated by the states. The purpose of initial teacher licensure was to ensure that all students benefitted from having qualified teachers

in the classroom. Obtaining teacher licensure grew increasingly complex because of the different requirements maintained by each state (Mitchell et al., 2001). States imposed numerous and varied requirements on prospective teacher candidates to better prepare them for the ultimate goal of obtaining licensure. States required candidates to complete their education, fulfill supervised practicum or internship requirements, pass required exams, provide evidence of good character, and meet other licensure requirements deemed necessary.

Traditional and Alternative Licensure Programs

Alternate routes to teacher certification provide prospective teacher candidates with the means of obtaining licensure, even if they were not an education major as an undergraduate student. At the same time, they provide opportunities for school divisions to recruit and hire teachers who have subject matter knowledge, but who pursued an alternate route to becoming a teacher. School districts provide these prospective teachers with on-the-job training, mentorship, and support leading to licensure (Feistritzer, 2008). Alternative licensure programs typically are designed to recruit, prepare, and license prospective teachers who already have a bachelor's degree, and who often have experience in careers other than education. In addition, they often require a rigorous screening process, such as passing tests, interviews, and competency in their subject matter. Alternative licensure programs permit prospective teachers to complete coursework and gain experience while being assigned as the teacher of record. These programs also require that licensure candidates work with a mentor (Feistritzer, 2008). The minimum grade point average (GPA) requirement for prospective alternative

certification candidates is usually 2.75, while traditional certification programs require that candidates have a GPA of 2.50 (Walsh & Jacobs, 2007).

One notable alternative certification program is called Teach for America. It was founded in 1989 at Princeton University in New Jersey, the first state to develop an alternative licensure program. Responding to a shortage of teachers, the New Jersey Department of Education sent recruiters to Ivy League schools to solicit Liberal Arts students to think about teaching (Feistritzer, 2008). This led to our current standard for alternative licensure programs, in which school divisions recruit prospective candidates for hard-to-fill teaching vacancies. The state of Texas adopted the Troops to Teachers recruitment program, in part because President George H. W. Bush had endorsed the program during his presidency. President Bush selected Rod Paige as the United States Secretary of Education because he had been Superintendent of the Houston Independent School District, which had the first and one of the largest district-run alternative licensure programs in the state (Feistritzer, 2008).

Traditional and alternative certification programs require candidates to complete a program designed ultimately to meet state certification requirements. Three distinguishing differences between traditional and alternative certification programs are: the controlling institution, the sequencing of certification requirements, and the terminology used to describe the teacher preparation programs (Evans, 2010). Traditional certification programs are housed in the schools or colleges of education on university campuses. Teacher candidates who choose the traditional route complete university-based courses with fieldwork requirements. These determine the focus of the preparation process and the systematic approach of becoming an expert in the content

area. Traditional programs typically require experience as the teacher of record and work with a mentor to ensure success during practicum. Teachers certified by traditional education programs must pass the state examination before becoming the teacher of record (Evans, 2010).

High-Stakes Testing

High-stakes testing is the process of attaching significant consequences to standardized test performance, with the goal of increasing teacher effectiveness and student achievement (Nichols, Glass, & Berliner, 2012; Herman & Haertel, 2005; Ryan, 2004). The rationale behind high-stakes testing is that teachers will be likely to work harder and students may be more motivated to learn by attaching incentives or threats to students' test scores. The repercussions for the teachers and the students include the following: penalties to the teachers and principals, school closure or "takeover", denial of diploma for students, and student grade retention.

Although the practice of high-stakes testing gained a prominent position in educational reform with the passage of NCLB, its use as a form of leverage preceded this mandate. The formal use of tests to distribute rewards and sanctions to teachers in urban schools began in the 1800s and has continued throughout the United States, especially since the 1970s (Berliner, Glass, & Nichols, 2012; Haertel & Herman; Tyack, 1974). In 1965, New York led the United States in test accountability efforts and implementing state-mandated minimum competency testing (MCT). New York also disseminated information to the media about local district performance on state assessments before NCLB (Berliner, Glass, & Nichols, 2012; Allington & McGill-Franzen, 1992). The enactment of NCLB mandated the most intrusive use of testing to influence how and

what teachers teach and how and what students learn. Despite an increase in research indicating that high-stakes testing has adverse effects on teaching practices and student motivation, policymakers continue to argue for its effectiveness to improve student learning (Berliner, Glass, & Nichols, 2012).

History of Student Achievement Testing

The Iowa Tests of Basic Skills (ITBS) has been administered over 60 million times to individual pupils in every state. The test has impacted instruction and teacher practices because of the data collected on student performance in specific areas. The ITBS, which was introduced in 1935, placed major emphasis on evaluating the development of basic skills rather than rote memorization. Grade-equivalent scales were used so that test results were comparable from test to test. Iowa norms reported on pupil and school achievement because of the availability of pupil profile charts and plotted reports of school averages (Lindquist, 2001). The ITBS also resulted in teachers being aware of those students who need individual help and those who were academically talented (Lindquist, 2001). Students began to be evaluated more frequently because of the ITBS. In 1940, the ITBS was extended to grades 3 through 5, and Houghton Mifflin Company published and distributed the tests. The new version of the ITBS test was called the Multi-Level Edition. This remains the current test that is used today (Lindquist, 2001). During his term as director of the ITBS program, Albert N. Hieronymus was responsible for revisions and the standardization of all subsequent forms (Lindquist, 2001).

No Child Left Behind

NCLB is the most far-reaching education policy initiative in the United States over the last four decades (Dee & Jacob, 2011). This legislation, which was signed by President Bush in January of 2002, changed the impact of federal influence over the nation's 90,000 public schools. This legislation made states accountable by making them conduct annual student assessments linked to state standards. NCLB includes annual testing in reading and mathematics in grades 3 through 8. It also includes testing once in grades 10 through 12 for reading and mathematics. The ratings of school performance – both overall and for subgroups – are factors in determining whether school districts are making adequate yearly progress (AYP) toward their state's proficiency goals (Dee & Jacob, 2011). NCLB also mandates that school divisions identify schools that fail to make adequate yearly progress. The goal of this legislation is to have all students proficient in reading and mathematics by 2013-2014 (Dee & Jacob, 2011). Sanctions and rewards are given based on each school's AYP status. Sanctions such as reconstitution of staff, public school choice, and school restructuring are mandated, especially for Title I schools that are persistently low-performing. Some states issue sanctions regardless of whether schools are considered Title I. The concept behind this reform is that linking students' academic performance to high-stakes tests will motivate school districts to monitor and improve the academic achievement of students in public schools. However, critics believe test-based accountability leads to educators who focus solely on tested subjects, shifting resources away from non-tested subjects (Dee & Jacob, 2011).

Race to the Top

Race to the Top (RTTT) was authorized under the American Recovery and Reinvestment Act of 2009 during President Obama's first term in office. RTTT required that school divisions interested in securing grants from the $4.35 billion fund use data to evaluate teacher performance and reward effective teachers (Hershberg & Robertson-Kraft, 2010). The new policy redefines important indicators used to measure student outcomes and teacher effectiveness. RTTT focuses on student growth over time and includes multiple measures for evaluating teachers. This information is used to determine compensation, tenure, and teacher advancement (Hershberg & Robertson-Kraft, 2010). The federal investment in RTTT is unprecedented and indicates that the federal government is ready for systemic change in education. The four educational reform assurances associated with RTTT are the following: rigorous standards and international benchmark assessments, the use of data systems linking student performance to teachers, recognition of great teachers and leaders, and turning around struggling schools (Hershberg & Robertson-Kraft, 2010). Like NCLB, RTTT emphasizes improving teacher quality as a step in the direction of improving student performance and closing achievement gaps. To receive funds, states must submit proposals that include student growth as one of the measures encompassing the teacher evaluation system. States must also propose plans to use this information in the decision-making process related to tenure, compensation, and promotions. States that refuse to use student data in teacher evaluations are not eligible to apply for funds (Hershberg & Robertson-Kraft, 2010). Delaware and Tennessee were the first two states to receive funds from RTTT in 2010. In its proposal, Tennessee stated that $12 million dollars in

competitive funding was available for school divisions willing to change compensation models to those that recruit talented prospective teachers (State of Tennessee, 2010). Offers of higher salaries to effective teachers and opportunities for advancement are also included in the new compensation model. Delaware changed their teacher evaluation system to allow for dismissal of teachers who demonstrate a pattern of ineffective performance over a two- or three-year period (Hershberg & Robertson-Kraft, 2010; State of Delaware, 2010).

Transition to Texas

Policymakers, politicians, administrators, and scholars have all voiced their opinions on NCLB. President Bush and the Secretary of Education Paige promised sweeping results from nationwide education reform (Nelson et al., 2007). While some positive changes in student achievement occurred in the years following the passage of NCLB, these changes have not been received without concerns from policymakers and stakeholders. Secretary of Education Paige found himself under scrutiny during his tenure as Superintendent of Houston Independent School District as a result of charges related to tampering with dropout statistics. He stated at the time that this was due to weakness among the district employees who did not follow his accountability system (Peabody, 2003). According to Nelson et al. (2007) exceptions were made throughout the state:

> Texas, widely considered the birthplace of NCLB, showed resistance and refused to comply with the rule, stating 1% of students with learning disabilities could be exempted from testing. The Texas Commissioner of Education gave acceptable accountability ratings to more than 900 schools

> with exemptions greater than 1%. This act was done to prevent the schools from being labeled low-performing because the school's population in this subgroup exceeded 1% (p. 702).

History of Texas Testing

The vision of education reform in Texas began in the 1990s and was led by the State Commissioner of Education, Kenneth H. Ashworth. The Texas model's conceptual framework was built on the belief that every student deserves a high-quality education regardless of the location of their school. Educational scholars believed student improvement in the aggregate group required improvement by only some of the students, rather than all students. The reform model of the Texas education system included attention to four major components: curriculum, assessments, accountability and a reporting system, and improved student learning (Nelson et al., 2007).

The state curriculum, called Essential Elements, was implemented during the 1984-85 school year. The purpose of the curriculum was to provide an instructional guide for teachers due to information gaps found in textbooks. The curriculum became the instructional standard for all school divisions in Texas. To ensure that teacher instruction was effective and that all students were learning, the state developed a statewide assessment test. The assessment tests were called the Texas Assessment of Basic Skills (TABS) and were used throughout the 1980s. However, the tests only measured mastery of the minimum skills. The new test was developed during the 1990s and called the Texas Assessment of Academic Skills (TAAS). It was developed as a criterion-referenced test and was aligned with the state curriculum. This test proved to be a diagnostic tool used to guide instruction. When the curriculum was again revised in

1996, the Texas Essential Knowledge and Skills (TEKS) was released simultaneously. However, this was not considered to be an effective measure of student knowledge. It consisted of two components: foundation and enrichment. This test focused on testing students only in the four primary content areas: reading, mathematics, social studies, and science. This test was later replaced by the Texas Assessments of Knowledge and Skills Tests (TAKS) in 2003
(Wong & Nicotera, 2007). TAKS was then used until the State of Texas Assessments of Academic Readiness (STAAR) was developed.

STAAR was initially administered to students beginning in the spring of 2012. Students in grades 3 through 8 would be tested in reading and mathematics. Students would take the writing test in grades 4 and 7, the science test in grades 5 and 8, and the social studies test in grade 8. To meet the federal requirements of NCLB, an alternative STAAR test was developed. The purpose of this test was to evaluate students in grades 3 through 8 who had specific learning disabilities and who received special education services (Texas Education Agency, 2013).

History of Texas Licensure

The Texas State Board for Educator Certification (SBEC) passed new laws in 1995 that included creating a renewable teacher license for Texas educators. Before this law, teachers were granted a lifetime teaching certificate in their content area. Prospective teacher candidates were now required to take a basic skills test in reading, writing, and mathematics before entering a teacher preparation program. This requirement began in 1989. Teacher candidates were required to have a bachelor's degree, coursework in a broad general education area, specialization in an academic area,

and teaching knowledge and skills. The examination for the Certification of Educators in Texas is given once the requirements for licensure are met (Cordova, 2001).

There are currently five ways that prospective teachers can obtain their teacher license in Texas. The first route is a traditional university-based program approved by the SBEC. This pathway usually results in a baccalaureate degree. The second method of licensure is an alternative program approved by the SBEC. These programs usually involve the completion of coursework, mentoring and supervision, and professional development experiences. The alternative programs are usually monitored through an education service center or school district (Cordova, 2001). A third route to licensure is certification by examination, in which the teacher candidate takes an exam and gains licensure by taking an exam in another content area and successfully passes the test. This route is for teachers who already have a Texas license. A fourth route involves the recognition of out-of-state certifications. The SBEC carefully evaluates all out-of-state certifications and determines whether or not to grant certification to the individual. Evaluations are based on comparable scores on the Praxis II subject area tests. The final route for certification is emergency certification. This is a right given to school divisions to hire personnel in critical areas established by the SBEC. These individuals must have five years of working experience with a teaching license (Texas Education Agency, 2013).

Problem Statement

This quantitative study sought to determine if alternatively-certified teachers and traditionally-certified teachers are equally equipped to prepare students to pass high-stakes tests in grades 3 through 8 in the state of Texas.

Since the inception of NCLB, teachers from numerous states have become concerned with the high-stakes testing component of this federal mandate and how student scores impact school accreditation. Specifically, they are concerned about the connections between poor test scores and sanctions if AYP is not met based on end-of-the-year summative assessments (Zimmerman & Dibenedetto, 2008).

Due to the increase in accountability associated with high-stakes testing, educators question the fairness of using student performance on standardized tests to evaluate instructional effectiveness (Zimmerman & Dibenedetto, 2008). Many teachers feel the need to "teach to the test", which may not be reflected in their curriculum requirements and instructional pacing guides. Many policymakers suggest eliminating the standardized test as a measure of teacher effectiveness in the classroom because of this pressure.

Purpose of Study

The intent of this quantitative study is to reveal whether there are any significant differences in the test scores on end-of-the-year state summative assessments when comparing the student scores in classrooms taught by traditionally-certified teachers and those taught by alternatively-certified teachers. This study aims to determine whether or not teacher licensure route has any effect on student test scores.

Significance of Study

The goal of this study is to provide insight into the extent to which traditional teaching programs and alternative certification programs equally prepare candidates with the pedagogy and instructional skills necessary to ensure that students are able to pass

high-stakes tests. This study can serve as a model for states seeking to link student performance to their teachers' route of certification. This study may also lead to an evaluation of how states and universities prepare prospective teachers for the classroom. It is clear that the policy positions of the Obama Administration support test-based accountability, and the government supports the policies' potential impact on the governing of our schools (Nichols, Glass, & Berliner, 2012).

Assumptions

The following are the assumptions made in the current study: a) the Texas Education Agency's database will screen out data for those students who take alternative tests due to special needs; b) the data will include years of teaching experience among teachers to decrease internal validity issues; c) the students' scores from the two teacher types will come from the same schools; d students who tested more than once on the same test are removed; e) the socioeconomic status of each students will be included; f) the current accreditation status of the schools will be included in the data; g the designation of the schools as Title I, urban, rural, or suburban will be included in the data; and h) the traditionally and alternatively-certified teachers will receive the same opportunities for professional development and guidance from instructional leaders when teaching the curriculum and planning instruction.

Overview of Methodology

This quantitative study will use an ex- post facto research design. This is appropriate because there is no direct manipulation of the independent variable. The "cause" has already occurred during the course of the experiment (Leedy & Ormrod,

2005). Ex- post facto designs are considered to be an example of passive observation designs. The researcher will use an ex- post facto design to examine relationships between naturally occurring population parameters and specific variables. The researcher will not manipulate the independent variable, but will examine it in relation to one or more variables for predictive reasons. This design will include a large number of subjects and will allow the researcher to examine phenomena that have already occurred. Random sampling, manipulation of the variables, and a control group are not present in this design (DePoy & Gitlin, 2011).

Research Questions

1. Are there significant differences between alternatively-certified and traditionally-certified teachers' scores on the following content area STAAR tests?
    - Grades 3-8 mathematics
    - Grades 3-8 reading
    - Grades 4 and 7 writing only
2. Is there a measurable difference between student scores in classrooms led by alternatively-certified teachers and student scores from traditionally-certified teachers in rural, suburban, and urban schools in Grades 3 through 8?

Organization of the Remainder of Study

Chapter II provides a review of the literature related to alternative certification teacher programs, traditional certification teacher programs, alternatively-certified teachers and student achievement, traditionally-certified teachers and student

achievement, the history of licensure, NCLB, and RTTP. Chapter III provides the research design for this study, including the population, sampling procedures, data collection process, instruments, and statistical methods used to answer research questions. Chapter IV details the analysis of the data and discussion of the findings. Chapter V contains the summary, conclusions, and recommendations of the study.

CHAPTER II

LITERATURE REVIEW

This literature review reveals the current research available regarding whether alternatively-certified teachers and traditionally-certified teachers' educational pathways impact student achievement. The review begins with a brief synopsis of teacher licensure, No Child Left Behind (NCLB), Race to the Top (RTTT), and a comparison of alternative and traditional certification teacher preparation programs. The focus of this literature review is to reveal the extent to which research indicates whether a teacher's certification pathway impacts student achievement on high-stakes testing. The theoretical framework upon which this study is built will be explained, followed by a discussion of the major studies exploring the research questions. The researcher has provided an examination of the studies that used quantitative analyses in support of this subject in order to support the validity of this study.

Licensure

The majority of states have a two- or three-tiered licensure process, with additional requirements to obtain a continuing or standard license. This process involves taking additional coursework, obtaining at least a bachelor's degree, and classroom experience. In all, 31 states require an initial license (which is valid for two to three years) and the attainment of a standard professional license based on the fulfillment of additional coursework. Another 13 states offer advanced licensure, and 3 states grant lifetime certificates at the advanced level (Knowles et al., 2001). Most states try to improve the quality of teaching by awarding licensure based on requirements in

connection with gaining additional experience, obtaining an advanced degree, and exceeding performance requirements in the classroom.

Due to the licensure requirements of the states, the supply of credentialed teachers has decreased in the content areas of mathematics, science, special education, and bilingual education in rural and urban areas. States have responded to teacher shortages in these critical areas by issuing various restricted licenses, allowing school divisions to hire teachers on a temporary or emergency basis. Some states have issued emergency licenses to individuals who possess a bachelor's degree, have passed a basic skills test, or hold a certification in another state. State rules differ as to which licensure requirements may be waived for teachers using emergency or temporary licenses. In 2006, 50 states and the District of Columbia reported that they were implementing a total of 125 alternative routes to teacher licensure to alleviate teacher shortages in the United States (Feistritzer, 2008).

No Child Left Behind Act (NCLB)

January 2002, NCLB was signed into law by President George W. Bush. The new law was a reauthorization of the Elementary and Secondary Education Act, which was originally enacted in 1965 as part of the Lyndon Johnson's War on Poverty (Rudalevige, 2003). It has been reauthorized every three to six years under a new name. Its main program initiative is Title I, which allocates nearly $12 billion dollars annually to schools in support of educating disadvantaged and underserved populations. NCLB required that states develop content and performance standards for K-12 schools. Congress also adopted the notion of Adequate Yearly Progress (AYP), which required states to make continuous progress toward the goal of academic proficiency for all

students. The goal was for all students to pass state tests by 2013-14. The mandates of NCLB to hold schools accountable and improve the educational experiences of all students have enlightened the public about student achievement gaps. Additional initiatives of NCLB include improving teaching quality and providing highly qualified teachers in the classrooms. The legislation considers new teachers to be highly qualified if they receive state certification and demonstrate content knowledge of the subject they teach, either by passing an subject-area exam or by having an undergraduate major in that subject or both (Boyd et al., 2007).

Race to the Top (RTTT)

The Obama Administration's RTTT competitive grant program has been recognized for revolutionizing the federal government's role in education and transforming state school reform efforts (McGuinn, 2011). RTTT differs from other federal programs because it supports only those states that have strong track records, plans for innovation, and a commitment to reform. RTTT is designed to use incentives instead of sanctions to drive state reform efforts. Part of the American Recovery and Reinvestment Act (ARRA) of 2009, RTTT emerged from $4.35 billion in funding that was allocated by Congress for state incentive grants. States must apply for participation and applications are graded on a 500-point scale according to specific criteria. These criteria include rigor of reforms proposed and their compliance with the four administrative priorities of RTTT. These priorities include: the development of common standards and assessments; improvements to teacher training, evaluation, and retention policies; the creation of useful data systems; and the adoption of school turn-around strategies (U.S. Department of Education, 2014).

All states applied during the first round of applications except Alaska, North Dakota, Texas, and Vermont. In January 2010, Delaware and Tennessee were awarded grants of $100 million and $500 million respectively (McGuinn, 2011). In all, 35 states and the District of Columbia applied during the second round of applications. Ten states were selected as recipients of the grants: Florida, Georgia, Hawaii, Maryland, Massachusetts, New York, North Carolina, Ohio, Rhode Island, Washington, and the District of Columbia. This program presents a new approach to federal education policy in the form of a competitive grant program. This program has generated substantive state policy changes in a short time span.

Comparing Traditional and Alternative Certification Teacher Programs

Traditional teacher preparation programs are the primary source of teacher supply in most states. These programs are regulated by the states. They are subjected to the criteria of accreditation groups and requirements mandated by individual programs and institutions of higher learning (Boyd et al., 2007). Prospective teacher candidates who successfully complete approved traditional preparation programs need only to pass the required state certification exams to become certified. States assume that, by completing the state-approved program, teachers have met the requirements for certification. This includes the required coursework, student teaching, practicum, and field experiences.

Researchers in New York have found that teacher certification from some traditional programs is a significant predictor of student achievement (Board of Regents, 2008; Darling-Hammond, 2010). A similar study of teachers in New York City also found that teachers' certification, graduation from a competitive college, and mathematics SAT scores were significant predictors of teacher effectiveness in elementary and middle

school mathematics. Student achievement was most enhanced by having a certified teacher who had graduated from a university pre-service program, had a strong academic background, and completed more than two years of teaching experience (Boyd et al., 2007; Darling-Hammond, 2010). The New York City team of researchers found that exemplary traditional teacher preparation programs have the following characteristics: they closely monitor the quality of student teaching experiences; they facilitate the match between the context of student teaching and candidates' teaching assignments; they ensure coursework in reading and mathematics; they help prospective teachers make practical application of skills learned in clinical experiences; and they incorporate a capstone project, which includes work completed with actual students (Darling-Hammond, 2010). These findings are similar to those provided by other researchers who similarly conducted studies of effective programs (Darling-Hammond, 2006; 2010; Zeichner, 1993). These researchers also concluded that an effective teacher preparation program typically has a strong clinical experience and curriculum for prospective teachers. Darling-Hammond (2010) stated that effective traditional programs teach candidates to analyze and apply what they learn in curriculum planning, use teaching and instructional strategies, and use performance assessments centered on the professional teaching standards when evaluating students. The most powerful traditional programs require students to spend time in the field throughout their program, examining and applying strategies while simultaneously while completing coursework. Traditional certification programs require candidates to work along with mentors and teachers who can show them how to engage learners in the classroom.

However, teaching candidates sometimes confront issues in both traditional and alternative programs. This may include the absence of the opportunity to receive direct modeling from expert teachers. Furthermore, recent research suggests prospective teachers must have opportunities to analyze various classroom situations in order to be productive. This may include situations that require candidates to think critically and reflect on their practices in and outside their student teaching or practicum experiences (Darling-Hammond, 2010).

In the United States, 20% to 30% of new teachers receive licensure from alternative certification programs (Kee, 2012; National Research Council, 2010). According to researchers, this proportion is larger than it was 20 years ago, when virtually all new teachers completed traditional certification programs and only a few thousand were alternatively certified (Feistritzer, 2007; Kee, 2012). These programs differ from traditional certification programs in many ways. They often seek to fill specific content area teacher shortages, and they enlarge the teacher applicant pool by recruiting diverse groups of prospective teachers (Kee, 2012). Alternative certification programs typically involve a period of intensive academic coursework and supervised, on-the-job training, in which prospective teachers learn the skills necessary to teach. Most alternative certification programs target mid-career switchers who already have earned a bachelor's degree (Blackburn et al., 2006). Candidates are normally eligible for a regular teaching license after a two- to three-year probationary period. Alternative programs differ according to quality, duration, and effectiveness. The programs can be national, state, local, or regional. Quality programs provide coursework that covers how

to effectively use instructional strategies to engage students. They also typically provide internships and mentorships by skilled professionals.

Three notable alternative certification programs are Troops to Teachers, Teach for America, and the New York City Teaching Fellows program. The Department of Defense established the Troops to Teachers program in 1994 to help improve public education by allocating funds to recruit and support former members of the military services as teachers in high-poverty areas (Blackburn et al., 2006; Troops to Teachers, 2004b). Congress passed the Troops to Teachers program in 1999 as part of Title XVII of the National Defense Authorization Act for Fiscal Year 2000. This was enacted to assist eligible members of the armed forces to obtain licensure as highly qualified teachers. This program supports prospective candidates to become elementary or secondary teachers. Eligible candidates include military retirees, members of active duty with an approved retirement date within one year of applying to the program, and honorably discharged service members with six or more years of service who are willing to be obligated to the Selected Reserves for three years.

New York City Teaching Fellows Program (NYCTF) was created in the summer of 2000 and has accounted for most of the growth in alternative certification in the city. The number of teaching fellows hired grew from 350 during the 2000-01 school year to 2,500 during the 2003-04 school year (Kane, Rockoff, & Staiger, 2008). This program was initiated following the 1999-2000 school year, when 60% of new teachers hired by the New York City Department of Education were uncertified. Due to the changes in New York State law that made certification requirements more rigorous, the New York

City Department of Education stopped hiring uncertified teachers and expanded its recruitment of alternative certified teachers.

Teach for America was formed from the blueprint of a college thesis of Princeton graduate Wendy Kopp, class of 1989. Kopp felt that top-notch college graduates might also feel a calling to teach. Teach for America (TFA) appeared to be similar to the old teacher corps. However, Kopp imagined the program to be selective, competitive, prestigious, and funded by corporate donations, grants, and federal funding (Feistritzer, 2008; Mabry, 1990). Teach for America members must meet specific requirements, make a two-year commitment, and demonstrate proficiency in the subject areas in which they will teach (Feistritzer, 2008). Teach for America employees work with school districts, states, and education programs to ensure that members are qualified to teach in the school divisions. Teach for America has developed partnerships with graduate schools so that teachers can pursue their Master's degree. Teach for America is now a public-private partnership with a 2011 operating budget of $309,115,182 dollars (Veltri, 2012). Corporations, foundations, and individuals fund 70% of the operating budget. In addition, 8% of public funding comes from the federal AmeriCorps program and the school districts in which the teachers are assigned (Feistritzer, 2008).

Urban teacher residencies (UTR) represent a fourth educational pathway that teachers can follow to obtain licensure. This program combines elements of traditional and alternative teacher preparation programs. Through a competitive process, uncertified teachers work with mentors for a year before becoming the teacher of record. Residents complete a streamlined set of coursework that leads to teacher certification and a Master's degree (Fullerton et. al, 2012). In exchange for tuition and a residency stipend,

prospective teachers commit to teaching in a school division for three to five years. The federal government has devoted more than $143 million since 2009 to establishing or expanding 26 residencies (Sawchuk, 2011). Proposals have been made to include additional funding for this program in the next reauthorization of the Elementary and Secondary Education Act (Berry et al., 2008; Papay et. al., 2012). The UTR model has spread rapidly since the first programs were launched in Chicago, Boston, and Denver between 2002 and 2004. Programs in New York, Philadelphia, and Los Angeles have met the Urban Teacher Residency support organizations' standards. Despite their growth, the UTR programs have not been formally evaluated or compared in studies to other newly hired teachers in terms of raising student achievement (Papay et al., 2012; National Academy of Education, 2008).

Theoretical Framework

Licensure is designed to guarantee a basic level of quality or skill of teachers in schools. There are several mechanisms through which state policy and criteria may affect the number of prospective teachers that end up teaching in the classroom (Brewer & Goldhaber, 2000). Brewer and Goldhaber (2000) stated that teachers' scores on their certification exams are predictors of teachers' performance in the classroom. Sawyer and Strauss (1986) used statewide data from North Carolina and found that district performance on standardized tests increased with the average teacher performance on the National Teacher's Exam. Ferguson (1998) found that teachers with higher than average performance on the state certification exam enjoyed similarly high student performance on the mathematics test. According to Brewer and Goldhaber (2002), there have been no

studies that used national data to examine the relationship between teacher licensure and student outcomes before 1988.

When determining whether a teacher candidate is ready to embark on their career as a classroom teacher, teacher preparation programs ensure that potential teachers are given numerous opportunities to practice and master concepts presented by their instructors and mentors in their certification programs (Brinkman, 2004). Collins (1988) described the Situated Learning Theory, which entails learning and coaching as definitive aspects of the apprenticeship style of learning. Situated Learning Theory incorporates real life scenarios, from everyday encounters to even the most theoretical occurrences, when beginning to prepare potential teacher candidates for the classroom (Brinkman, 2004). Brill (2001) further described the Situated Learning Theory as a process that encourages teachers to immerse their students in an environment that cultivates appropriate behaviors and facilitates the mastery of strategies needed to become an effective classroom teacher.

Meyer (2011) described Situated Learning Theory as "on-the-job training with invaluable benefits for students in teacher preparation programs" (p. 143). This involves collaboration, teamwork, leadership, reflection, critical thinking, and shared decision-making. According to Meyer (2011), the Situated Learning model is becoming the blueprint for several teacher preparation programs. This model encompasses the student teaching component of many teacher certification programs. It also allows potential teacher candidates to learn, apply, and practice effective instructional strategies in a real - life setting. Training teachers in this model gives potential teacher candidates

opportunities to practice teaching and learning in the apprentice-style model, which provides them with experiences which they can later emulate in their own classrooms.

Teacher Preparation and Student Achievement

Several studies since 2000 have been conducted to determine if there is a relationship between student achievement and teacher certification pathway. The findings of these studies have concluded that teachers who received certification from traditional programs have a positive impact on student achievement as compared to teachers earning licensure from non-traditional programs. However, others connect alternatively-certified teachers with student performance that is equal or superior to the student achievement in classrooms taught by teachers who attended traditional programs from colleges and universities. The purpose of these studies is to determine the impacts of licensure pathway on student achievement. These studies are generally divided into two main categories: those reporting that the students of traditionally-certified teachers enjoyed higher levels of achievement, and those reporting that students of alternatively-certified teachers achieved comparable or superior scores to students taught by traditionally-certified teachers. The following is a summary of studies in both categories.

Benefits of Traditional Certification on Student Achievement

Goldhaber and Brewer (2000) conducted a longitudinal study determining whether teacher certification mattered when comparing the student scores of traditionally and alternatively-certified teachers from 1988-90. This longitudinal study compared teachers with standard certification to those with emergency licensure, probationary licensure, and teachers who taught out of their content area. The study linked teachers to

student data using the National Educational Longitudinal Study of 1988. Students were tested in high school mathematics and science. The results indicated that teachers with standard certification in mathematics and those who were certified and taught out of their content area had students who achieved a 1.3-point increase on the mathematics test. This is equivalent to 10% of the standard deviation on the 12th grade test. The results for the science tests were identical. Teachers who were not certified ultimately had a negative impact on student science test scores. The students of teachers who were emergency-certified scored the same as the students of the traditionally-certified teachers. One limitation of this study was that the data set did not distinguish between teachers without certification and certification out of the subject area. Brewer and Goldhaber (2000) concluded that the abilities of teachers certified on an emergency basis were comparable to those of traditionally-certified teachers because they went through a screening process for content knowledge. However, further studies would be needed to substantiate this claim. In turn, Berry, Darling-Hammond, and Thoreson (2001) criticized the methodology of this study regarding sampling and multicolinearity.

Research by Berliner and Lackzo (2002) and Hammond et al. (2005) followed longitudinal studies to determine if the students of traditionally-certified teachers performed better than those of alternatively-certified teachers. Berliner and Lackzo-Kerr (2002) conducted a longitudinal study in 1998-2000 and linked students to teachers according to their certification pathway; whether they were traditional or alternative. This study took place in Arizona and data were collected on teachers who taught grades 3 through 8 in five elementary schools. The instrument used to compare student scores was the Standard Achievement Test, Ninth Edition (SAT9). The SAT9 was believed by the

State Department of Education to relate specifically to Arizona's academic standards, which teachers used to drive instruction. Under-certified teachers in this study were classified as emergency, temporary, or provisional. A correlated *t*-test was used to evaluate whether there is a difference in student achievement scores among certified and under-certified teachers. Results indicated that students taught by certified teachers outperformed students taught by under-certified teachers in reading and language arts. Mathematics scores did not show a significant difference. During the 1999-2000 school year, results were the same as the previous year for reading, language arts, and mathematics. According to Berliner and Lackzo-Kerr (2002), the academic year was 10 months long, so the loss of two months, or $1/6^{th}$ of the year, was incurred by students in classrooms taught by under-certified teachers. When comparing the scores of Teach for America teachers with the scores of students in classrooms taught by teachers who were considered un-certified, Teach for America student scores were not significantly different in mathematics, reading, and language arts (Lackzo-Kerr & Berliner, 2002). On all tests from 1999-2000, certified teachers outperformed under-certified Teach for America teachers based on student scores. The results from this study contradict the claims made by Teach for America advocates that their education from some of the most prestigious universities and training prepares them for teaching students (Lackzo-Kerr & Berliner, 2002). However, one limitation of this study was the lack of controls used for prior student achievement, which was not considered by the researchers.

Hammond et al. (2005) performed a similar longitudinal study from 1995-2002, which linked teacher certifications to student achievement. The research used student data from a Houston Independent School Division using test results from three high-

stakes assessments: TAAS, SAT, and the Aprenda, the Spanish language test (Hammond, et al., 2005). This study was similar to a previous study conducted for the Hoover Institution's CREDO Center (Raymond, Fletcher, & Luque, 2001). The study examined the effects of Teach for America teachers on student achievement gains in mathematics and reading from grades 3 through 8 using the TAAS test from 1996-2000 (Hammond et al., 2005; Raymond et al., 2001). The results of this study revealed that the students of Teach for America teachers performed comparably to teachers with one or zero years of teaching experience, after controlling for teacher experience. Contributing factors that differ from the study by Hammond et al. (2005) included the fact that the Teach for America teachers were not compared to traditionally-certified teachers in the Houston Independent School District. Rather, Teach for America teachers were compared to uncertified teachers because during the 1999-2000 school year (the last year of the study) 50% of all new teachers and one-third of all teachers in the district were uncertified. Researchers reported that many teachers lacked a bachelor's degree (Hammond et al., 2005). In addition, a longitudinal study by Hammond et al. (2005) included student data from grades 3 through 8 and grade 10. The student scores collected from the criterion-reference test (TAAS) consisted of two metrics: whether or not the student met minimum academic expectations and the Texas Learning Index (TLI), which allowed comparisons of student learning progress between grades. The SAT9 was introduced in 1997-1998 to include a nationwide standardized test for students in Grades 1 through 11 for reading and language arts. In addition, Aprenda was administered to students in Grades 1 through 9 who received instruction in Spanish. During this research

study, the following variables were controlled: prior student achievement, student demographics, and teachers' years of experience.

The Houston Independent School District conducted their own alternative certification program in which Teach for America teachers were placed in the school division's program upon being hired. Many of the Teach for America teachers were certified by their 2nd or 3rd year of teaching (Hammond et al., 2005). Teach for America had a positive effect on student achievement in mathematics and an insignificant effect in reading, as evaluated by the TAAS test. On the SAT9 and the Aprenda assessments, Teach for America teachers had a negative effect on student scores in mathematics and reading. On the TAAS reading test, which the Teach for America coefficient in the pooled years' analysis had been insignificant, the Teach for America coefficients were significant and positive in 1998-99. This was largely because this was the school year in which Teach for America recruits were more likely to be certified than most Houston teachers. In this year, 73% of the Teach for America teachers had earned a traditional certification. By comparison, only 65% of the remaining Houston teachers had a traditional certification. During the 1999-2000 school year, only 46% of Teach for America teachers had their certification, in comparison to 68% of the other Houston teachers. This correlates to the shift in student scores among Teach for America recruits, which demonstrated an insignificant or negative impact on each test.

This study revealed that Teach for America and Houston's Independent School Division's alternative teacher certification program had negative effects on achievement on three tests. In addition, teachers who held emergency or temporary licenses showed a negative effect on three tests. However, they showed a positive effect on the SAT9

Reading scores. Certified teachers teaching out of their content area showed positive effects on the following tests: the reading and mathematics TAAS tests, the Aprenda test in reading, and a positive coefficient in reading on the SAT9 (Hammond et al., 2005). Teachers without certification or with non-standard certification were found to be less effective in raising student test scores than teachers with standard certification in 22 of 36 estimates ($p<10$). When comparing teachers with standard certification to teachers with nontraditional certification, student progress declined over the course of a year by about 1/2 to 1 month in grade equivalent terms on most achievement tests. When comparing the studies by Hammond et al. (2005) and Berliner and Lackzo-Kerr (2002), the similarities among the findings revealed the following three trends: a) teachers who held traditional certification were required to have passed certification exams in the subject matter they intended to teach, b) teachers were taught how to design and implement instruction, and c) teachers were taught how to communicate effectively with parents and students. Teachers also completed an approved traditional teacher preparation program that included coursework specific to their content area. This coursework showed teacher candidates how to engage students and taught them how to effectively use instructional strategies and evaluate students (Hammond et al., 2005; Texas Administrative Code, Title 19, Part 7, Rule 230.191, 2004).

Clotfelter, Ladd, and Vigdor (2007) conducted a study in North Carolina comparing teacher credentials and student achievement in high school. The study compared teachers with regular or continuing licensure who completed a state-wide approved teacher preparation program with teachers who are considered lateral entry certified. The laterally certified teachers are considered alternatively-certified because

they have a bachelor's degree, but they have not completed their coursework (Clotfelter, Ladd, & Vigdor, 2007). Laterally certified teachers must be affiliated with a college or university and complete six hours of coursework each year. The study revealed that the achievement of students who had a teacher with a lateral license was reduced .06 standard deviations when compared to the scores of students who had a teacher with traditional certification (Clotfelter et al., 2007). The results indicated that being certified in a specific content area is predictive of high student achievement. Laterally certified teachers with some teaching experience appeared to be no less effective than teachers with a traditional license. This is possibly attributed to the fact that many teachers had some on-the-job training during their first two years of teaching. This may also be attributable to the selection process. Similarly, this study found that alternatively certified teachers performed worse than traditionally licensed teachers in high school content area (Clotfelter, Ladd & Vigdor, 2007).

Positive Impacts of Alternative Certification

A random assignment study of student achievement of traditionally-certified and alternatively-certified teachers by Constantine et al. (2009) indicated that there were no significant differences in student achievement between them. Croninger et al. (2003) similarly found no difference in teacher effectiveness for first grade students from different certification routes.

Blackstone (2010) performed a quantitative correlation study of teacher preparation program effect on student achievement in rural school settings. The study compared test scores of traditionally-certified and alternatively-certified teachers' student scores on grades 7 and 8 end-of-course mathematics test. This study revealed no

significant difference of measured effect on student achievement among traditionally and alternatively trained teachers.

Teach for America is considered to be one of the notable alternative teacher preparation programs in the country. Decker, Glazerman, and Mayer (2006) conducted a study comparing student scores associated with Teach for America teachers to traditionally-certified teachers, other alternatively-certified teachers, and teachers considered uncertified. The subjects tested included reading and mathematics, using a pre-test and post-test of the Iowa Test of Basic Skills (ITBS). The sample of Teach for America teachers came from the 2000-2002 cohorts. All of the Teach for America teachers in this study had master's degrees, while teachers from the other groups all had bachelor's degrees. The results of the study revealed that the students from Teach for America teachers excelled more than students in the uncertified and certified teachers' classes in mathematics. The Teach for America students increased their ranking from the 14^{th} percentile to 17^{th} percentile, while the other students remained in the 15^{th} percentile in the fall and at the end of the year in mathematics. The overall impact of Teach for America teachers on student performance was roughly a 10 percent grade equivalent, which corresponds to an additional month of instruction. The impact on the reading scores of the Teach for America students was not statistically significant, thought it was close to zero, or an effect size of 0.03. When comparing Teach for America teachers to all certified teachers, the mathematics impact was not as significant when compared to certified teachers. A possible reason for the decrease in scores for the Teach for America teachers is the lack of teaching experience (Glazerman et al., 2006).

Hannaway, Taylor, and Xu (2011) were the first to estimate the impact of Teach for America teachers on high school achievement. Using data from North Carolina, they found that Teach for America teachers improved student performance in mathematics, science, and English. These results were still valid when comparing student achievement among Teach for America students to student achievement among teachers with more years of teaching experience.

An alternative certification program, called Career Switchers, which was funded by a five-year grant from 2002-2007 by the U.S. Department of Education to the state of Virginia, resulted in a study that focused on whether student achievement was positively impact by alternative licensure. In this case, researchers compared the mathematics scores of students taught by Career Switchers to the scores of students in classes taught by traditionally-certified teachers (Bol, Gimbert & Wallace, 2007). Old Bay Public Schools and New Division University (OBPS-NDU), the recipient of the grant, formed a partnership to train teachers who had a bachelor's degree in another field to enter the Transition to Teaching Program (TTT). The higher education partnership during 2002-07 was established with the Darden College of Education. The selection process for teachers was highly selective and candidates had to have a passing score on the Praxis I and Praxis II tests in mathematics. A five-week summer institute was required for teachers, which focused on education coursework. This was followed by the requirement that prospective teachers obtain their professional license in secondary mathematics. In addition, each participant had to commit to a three-year assignment in an urban school district. A cohort of 1st year alternatively trained Algebra I teachers and 1st year traditionally trained Algebra I teachers were selected as participants. ANOVA was used to compare students'

Standards of Learning (SOL) scores in Algebra I based on the certification pathway chosen by teachers.

The results of the study revealed that the type of preparation program did not significantly influence the overall SOL test scores for Algebra I (Bol, Gimbert, &Wallace 2007). The results of the statistical test MANOVA on the district quarterly tests revealed statistically significant differences on the first and second district quarterly tests when comparing the student scores of alternatively-certified teachers to the scores of traditionally-certified teachers. In both cases, the students of traditionally-certified teachers scored higher. On the third district quarterly test, there were no significant differences between the student scores among the two teacher training programs. These findings suggested that teacher training had a significant influence on Algebra I student achievement before the final administration of the district quarterly tests. However, there was no significant influence on the students' Algebra I scores on the final district quarterly test. The results of study indicated that alternatively-certified teachers were capable of promoting positive student achievement in mathematics.

Kane, Rockoff and Staiger (2008) conducted a study that focused on teachers of mathematics and reading in Grades 4 through 8 from 1998-99 to 2004-05. During this time, New York City hired more than 50,000 teachers. Among them, 46% were certified, 34% were uncertified, and 20% were alternatively-certified. The majority of the alternatively-certified teachers were recruited from the New York City Teaching Fellows Program, while most of the others came from Teach for America. The math scores of the students assigned to the New York City Teaching Fellows Program and Teach for America students' scores were .20 and .28 standard deviations below regularly-certified

teachers' scores, respectively (Kane, Rockoff, & Staiger, 2007; 2008). All of the participants were 1st year teachers. The reading scores were similarly significant. These scores contributed to apparent differences in prior year test performance and demographics. When the test scores from the prior year were included, students of teachers from the New York City Teaching Fellows classes performed no differently than similar students assigned to traditionally-certified teachers in mathematics. Students assigned to Teach for America teachers outperformed traditionally-certified teachers by .01 standard deviations. New York City Teaching Fellows and Teach for America students underperformed as compared to students of traditionally-certified teachers in reading. During this study, the researchers found that negative effects were reduced or eliminated in mathematics as teachers finished their training and certification, and gained critical teaching experience (Heilig & Jez, 2010). Teach for America teachers continued to have a negative effect on reading throughout the duration of the study (Heilig & Jez, 2010).

Nunnery, Kaplan, Owings, and Pribesh (2010) performed a study in Florida comparing teachers funded by Troops to Teachers with traditionally-certified teachers. The study compared student performance between the two teacher types on the 2003 and 2004 Florida Comprehensive Assessment Test (FCAT) in reading and mathematics. Teachers' years of experience were matched in this study to eliminate selection bias. The study revealed that the students of Troops to Teachers outperformed students of traditionally-certified teachers. This study confirmed that the students of alternatively-certified teachers can perform comparably or better than the students of traditionally-certified teachers on standardized tests.

Another recognizable alternative certification program is the American Board Certification of Teacher Excellence (ABCTE). ABCTE teachers are certified in several states. Studies on the impacts of this program on student achievement have only been conducted in Florida, and the results so far are mixed. Using a matching method with 30 ABCTE teachers of students in grades 4 through 10 in Florida, Anderson, Clark-Tuttle, and Glazerman (2009) found that the student achievement in ABCTE teachers' classrooms was comparable to that of traditionally-certified teachers in reading. However, it was significantly worse in mathematics. These findings contrasted those of the study performed by Sass (2011), who found that ABCTE teachers' student performance exceeded that of traditionally-certified teachers in reading and mathematics.

Burns, Gansle, and Noell (2012) found mixed results when performing a hierarchical linear modeling study of student achievement among traditionally-certified and alternatively-certified teachers in Louisiana schools. The study revealed that the Private Practitioner Traditional Program and University Practitioner Program II students in grades 4 through 9 scored higher than the scores of students in the classrooms of average new teachers on the ITBS. The alternative teacher program – called the Masters Alternate Certification Program I – was the only alternative program that scored comparably to the scores of traditional programs.

Goldhaber, Liddle, and Theobald (2013) used a methodology that allowed teacher training effects to decay. The researchers evaluated whether or not teacher preparation had an effect on student achievement. They concluded that regardless of when teachers receive their certification or from what program, training program affects decay as teachers gain workforce experience. In other words, after three to five years of classroom

experience, there is no difference in student performance or retention between alternatively-licensed teachers and traditionally-licensed teachers. This is indicative of the fact that teacher training should not be thought of as invariant to a teacher's workforce experience (Goldhaber, Liddle & Theobald, 2013).

Boyd et al. (2009) conducted a study involving the evaluation of student performance in mathematics and language arts among first-year teachers in New York City from various teacher preparation programs. The study revealed differences across teacher preparation programs, which result in some teachers having a significantly greater effect on student achievement. For example, teacher preparation programs that have teachers who are effective in mathematics also produce teachers who are also effective in language arts. The actual features of teacher preparation programs can have a direct effect on student achievement as a result of the teachers' experiences and training. A program's ability to attract good candidates and the program's ability to provide direct training related to the classroom improve teachers' effectiveness. This pertains to all teachers, whether they are alternatively or traditionally-certified.

Summary

The studies available on whether teacher certification routes affect student achievement are ultimately inconclusive. More studies are needed using longitudinal data that link teachers directly to students. These studies should include teachers' years of experience, certification routes, and student demographics. To determine a definitive answer on whether teachers' certification routes affect students' achievement, the state departments of education need to be active participants in collecting and disseminating data. This will allow researchers to have the information needed to perform quantitative

studies. The results of the studies performed in the future will be valuable when determining the effectiveness of teachers in the classroom and improving student achievement on various types of assessments.

CHAPTER III

METHODOLOGY

The methodology described in this section was used to determine whether or not there is a difference between traditionally-certified teachers' student scores and alternatively-certified teachers' student scores on the Texas Assessments of Academic Readiness (STAAR). The entire universe of the sample was used to determine whether or not there was a significant difference in the students' performance on these high-stakes tests. The selection of the participants came from the Texas Education Agency, which links teachers' certification routes to students' performance on summative assessments such as the STARR tests. Student data were collected using data from the 2011-12 school year. A baseline was established by using 2011-12 data, which controlled for prior achievement when using data from 2012-13 to determine if significant gains in achievement occurred in the same content areas tested in 2011-12. The analysis used was the Analysis of Covariance (ANCOVA). This was used to determine if there are any differences in the group means of students' scores from alternatively- and traditionally-certified teachers. The other analysis used was the Analysis of Variance (ANOVA), which was used to determine if there were differences in the group means of traditionally-certified and alternatively-certified students' scores in rural, suburban, and urban areas.

Purpose of Study

The intent of this quantitative study is to reveal whether there are significant differences in test scores on the State of Texas Assessments of Academic Readiness (STAAR), an end-of-year summative assessment, when comparing students of

traditionally-certified teachers with those of alternatively-certified teachers. Ultimately, this study will determine if teacher licensure route has an effect on student test scores as well as on students from urban, rural, and suburban areas.

Research Design

The researcher used an ex- post facto design in this quantitative study. This was the appropriate design because there was no direct manipulation of the independent variable (Leedy & Ormrod, 2005). It was a causal comparative study. This design was used because it attempted to determine cause for existing conditions or pre-existing differences in groups. In this case, the cause and effect have already occurred. Ex- post facto was the selected research design because it attempted to identify cause-effect relationships by involving two or more groups in the study. Individuals will already be in groups before the study begins. The sample was not randomly selected.

Ex -post facto design was the most appropriate design for the study because the Texas Education Agency sampled from the population of alternatively-certified and traditionally-certified teachers who have taught in tested content areas from the grade bands of 3rd through 8th grades. The design was suitable because the direct manipulation and collection of the data were done by the Texas Education Agency's accountability department, which resulted in a valid sample of the population of alternatively- and traditionally-certified teachers who teach students in 3rd through 8th grades. The independent variable was not manipulated by the researcher, but by the accountability department of the Texas Education Agency. This ensured the validity of the collection of the data.

The strengths of this study included an examination of groups who were comparable to one another, a large sample size, and multiple replications of the treatment effect. In an ex- post facto research design, the researcher takes groups of individuals who are already different on some measurable variable and treats them all in the same manner before measuring them based on another variable (Sprinthall, 2012). This study also included determining whether differences in urban, rural, and suburban student scores existed when taught by traditionally- and alternatively-certified teachers.

Research Questions

1. Are there significant differences between alternatively-certified and traditionally-certified teachers' student scores on the following content area STAAR tests?
    - Grades 3-8 mathematics
    - Grades 3-8 reading
    - Grades 4 and 7 writing only
2. Is there a measurable difference between student scores in classrooms taught by alternatively-certified teachers and student scores from traditionally-certified teachers in rural, suburban, and urban schools in grades 3 through 8?

Participants

The selection of participants came from the Texas Education Agency, which identifies teachers as alternatively- or traditionally-certified in their database. These teachers were selected because their students' performance on the end-of-year summative assessments was linked to teachers based on their certification type (M. Ramsey, personal

communication, Jan. 24, 2013). Texas began linking students' performance on the STARR tests to teacher certification routes. (M. Ramsey, personal communication, Jan. 24, 2013). This was the reason for the selection of the Texas Education Agency for the study.

The assignment of participants and students included a numerical descriptor to protect their identity. The data request to the Texas Education Agency took one month to process (M. Ramsey, personal communication, Jan. 24, 2013). The data were already correlated with the teacher certification routes and students' scores.

Sampling

The participants consisted of alternatively-certified teachers and traditionally-certified Texas teachers. The sample consisted of the entire universe, which was the total population of alternatively- and traditionally-certified teachers who taught grades 3 through 8 throughout the state of Texas. The students were those in the classrooms of the alternatively- and traditionally-certified teachers. The student and teacher data came from the same schools. The teacher selection was performed by the testing and accountability department (M. Shim, personal communication, Jan. 23, 2013).

Instrumentation

The collection and the analysis of data occurred using the Statistical Package for the Social Sciences (SPSS). This statistical software program had the capability of tabulating and comparing students' scores from the summative tests and identifying existing trends when comparing students' scores among the two teacher groups. Analysis of Covariance (ANCOVA) was used to determine if there were any differences in the

group means of students' scores from alternatively- and traditionally-certified teachers. The one-way Analysis of Variance (ANOVA) was used to determine if there were significant differences in the group means of traditionally-certified and alternatively-certified teachers' student scores in rural, suburban, and urban areas. The independent variable (IV) was the communities in which the schools were located, and the dependent variables were the math scale scores, writing scaled scores, and reading scaled scores of the students.

Procedure for Data Collection

The selection of the participants in the study occurred by requesting data from the Texas Education Agency that linked students' scores on summative assessments of students in grades 3 -8 for mathematics, grades 3-8 for reading, and grades 4 and 7 writing. An official request for data was made to include the students' scores of alternatively- and traditionally-certified teachers who taught in urban schools, suburban schools, and rural schools within the same school division. The rationale was to determine if traditionally- and alternatively- certified teachers from schools with different socioeconomic status students' scores differed based on their teachers' certification routes. Numerical descriptors were requested by the researcher to conceal the identification of the students in the study. Control for prior achievement was accomplished by collecting student scores on the STAAR tests for 2011-12. Data were collected on the same students using 2012-13 data to determine if significant gains in achievement occurred over the year in the same content areas that were tested in 2011-12. The rationale for establishing a baseline was to reveal whether the students' assignment in traditionally and alternatively- certified teachers made an impact on students'

performance on the STAAR tests (A. Gallegos, personal communication, June 26, 2014). The data obtained from the Texas Education Agency consisted of all alternatively- and traditionally- certified teachers. Once the data were obtained, the data were analyzed using the Analysis of Variance (ANOVA) to determine whether any differences exist in the mean of the scores of students in Grades 3 through 8 among the traditionally- and alternatively-certified teachers from urban schools, rural schools, and suburban schools. Data analysis also included comparisons of the students' scores of the two teacher types and determinations regarding whether or not there is a difference in performance based on their teachers' preparation pathway using ANCOVA.

One-way Analysis of Variance (ANOVA) was used to determine if there were significant differences in the group means of traditionally-certified and alternatively-certified students' scores in rural, suburban, and urban areas. The independent variable (IV) was the communities in which the schools were located, and the dependent variables were the math scale scores, writing scaled scores, and reading scaled scores of the students.

Data Analysis

The sample size consisted of the entire population of the universe from the Texas Education Agency. Controlling for prior achievement was accomplished by establishing a baseline using students' performance on the 2011-12 STAAR tests. Data were collected on the same students using 2012-13 data to determine if significant gains in achievement occurred over the year in the same content areas that were tested in 2011-2012. The rationale for establishing a baseline was to reveal whether the students' assignment in

traditionally and alternatively- certified teachers made an impact on students' performance on the STAAR tests (A. Gallegos, personal communication, June 26, 2014).

Teachers of the students were identified as traditionally-certified or alternatively-certified prior to the administering of the tests. Students' teacher types were tracked during the administration of the 2012-2013 STAAR tests to determine whether students' performance increased, decreased, or remained the same when assigned alternatively-certified or traditionally- certified teachers during the 2012-2013 administering of the STAAR tests. Data were analyzed to determine if any trends existed when evaluating scores from students who had traditionally-certified or alternatively-certified teachers consecutively during the 2011-12 and 2012-13 school years.

For the students who were not assigned consecutive teachers from the same educational pathway from the 2011-12 and 2012-13 school years, academic growth was statistically measured to determine whether or not there was a significant difference in the students' performance when analyzing the data from 2011-12 to 2012-13 academic school years. The statistic analysis used in this study is the Analysis of Covariance. ANCOVA was used to evaluate whether or not the population means of a dependent variable (DV) were equal across levels of a categorical independent variable (IV), while controlling for the effects of other continuous variables such as covariates. When performing ANCOVA, the (DV) means were adjusted to what they would be if all groups were equal on the covariate. Covariance is a measure of how much two variables change together and how strong a relationship exists between them. Covariates can be a cofounding variable that can influence the (DV). The covariate in this study was controlling for prior achievement. ANCOVA was used in this study to determine if

significant differences between two groups existed by reducing the within-group variance. Another use of ANCOVA was to correct for initial group differences (prior to group assignment) that existed on (DV) among groups. In this situation, participants cannot be made equal through random assignment, so the covariate was used to adjust scores (Field, 2011).

One-way Analysis of Variance (ANOVA) was used to determine if there were significant differences in the group means of traditionally-certified and alternatively-certified students' scores in rural, suburban, and urban areas. The independent variable (IV) was the communities in which schools were located, and the dependent variables (DV) were the math scale scores, writing scaled scores, and reading scaled scores of the students.

Limitations

Limitations of the study are the following: a) the movement of students within the district, and the potential effect that poor record-keeping may have in jeopardizing the validity of the data; b) school divisions that provide instructional assistance to teachers whose students perform poorly on formative assessments and not to other teachers, which can cause internal validity issues; c) teachers from different schools within the school division may use different instructional strategies in their classrooms, which can impact students' scores.

Summary

In this study, the researcher examined the impact of teacher certification route on student achievement on the Texas end-of-year STAAR exam. The study involved the use

of quantitative research methods. Data from the STAAR test were analyzed to determine whether there was a significant difference of traditionally-certified teachers and alternatively-certified teachers' student scores on high-stakes standardized tests when controlling for prior achievement. The results of this study were used to reveal whether or not states need to examine the requirements for certifying teachers to teach in our nation's classrooms.

CHAPTER IV

RESULTS

The purpose of this study was to measure potential differences between traditionally-certified and alternatively-certified student scores on the STAAR tests in reading, writing, and mathematics. Data were collected for over a span of two years. Previous literature examined whether teacher years of experience and certification route affected student performance. The state of Texas was selected for this study because it links teachers to students' performance based on their certification route.

The research questions selected were the following:

1. Are there significant differences between alternatively-certified and traditionally-certified teachers' student scores on the following content area STAAR tests?

- Grades 3-8 mathematics
- Grades 3-8 reading
- Grades 4 and 7 writing only

2. Is there a measurable difference between student scores in classrooms taught by alternatively-certified teachers and the scores of students taught by traditionally-certified teachers in rural, suburban, and urban schools in grades 3-8.

The study presented whether statistically significant differences exist in the student scores of traditionally and alternatively certified teachers in 3rd-8th grade reading, mathematics, and writing. Results for 4th and 7th grade writing student scores of traditionally and alternatively-certified teachers were also presented in the study. By examining the results from the study, the researcher was able to determine whether

significant differences exist in the student scores of traditionally and alternatively-certified teachers on high-stakes tests such as the STAAR tests.

FINDINGS

Results for Research Question # 1

This chapter presents the findings of this study. Students' performance by grade level in mathematics, reading, and writing was analyzed using the ANCOVA statistical test, which determined if there were any significant differences in students' scores according to teachers' certification types. ANCOVA evaluated whether the population means of a dependent variable (DV) were equal across levels of a categorical independent variable IV while controlling for the effects of a continuous variable, a covariate. When performing ANCOVA, the DV means were adjusted to what they would be if all groups were equal on the covariate. Covariance is a measure of how much two variables change together and how strong a relationship exists between them. Covariates can be a cofounding variable that can influence the DV. The covariate in this study was controlling for prior achievement.

Descriptive Statistics for Writing

The descriptive statistic results shown in Table 1, indicated that more alternatively-certified teachers taught 4th grade writing, (N =703) while in 7th grade more traditionally-certified teachers taught writing, (N =508).

Table 1

ANCOVA of Between-Subjects Factors of Certification Types of 4th and 7th grade Writing

Between-Subjects Factors

Grade Level 2013			Value Label	N
4	Certification Type	1	Alternative Program	275
		2	Alternative Program (Post-Bac)	28
		4	Out-of-State	35
		5	Traditional Program	395
7	Certification Type	1	Alternative Program	430
		2	Alternative Program (Post-Bac)	91
		4	Out-of-State	55
		5	Traditional Program	508

As shown in Table 2, 7th grade writing scores for students in alternatively-certified teachers' classrooms and the writing scores of 7th grade students in out-of-state classrooms showed more variance than the test scores for other teacher certification

types. The variance gives information about how scores differ or vary. The standard deviation that is being compared in Table 2 indicates how scores vary from the mean. The larger the standard deviation value, the more the scores are spread out around the mean. The numbers of out-of state certified teachers were only 35 for 4th grade and 55 for 7th grade. The traditionally-certified teachers totaled 395 for 4th grade and 508 for 7th grade. The smaller number of out of state teachers may explain why the variance was greater.

Table 2

Descriptive Statistics of Writing Scale Score of 2013

Grade Level 2013	Certification Type	*M*	*SD*	*N*
4	Alternative Program	3454.66	398.447	275
	Alternative Program (Post-Bac)	3329.93	390.822	28
	Out-of-State	3440.69	515.470	35
	Traditional Program	3431.44	436.122	395
	Total	3436.71	424.679	733
7	Alternative Program	3294.91	549.665	430
	Alternative Program (Post-Bac)	3295.19	672.947	91
	Out-of-State	3307.89	692.739	55
	Traditional Program	3296.00	551.876	508
	Total	3296.10	569.062	1084

Table 3

Levene's Test for Equality of Variances for Writing Scaled Score

Grade Level 2013	F	df1	df2	Sig.
4	1.47	3	729	.220
7	.813	3	1080	.487

*p<0.05

Levene's Test of Equality of Variances Results

The first step before performing the ANCOVA analysis was to conduct the Levene's Test of Equality of Variances. The assumption was that there were no variances in the group means. The Levene's Test for Equality of Variances was used to test the assumption of homogeneity of variance. When performing the analysis, the level of significance was set at $\alpha = .05$. A significant result indicates that the variance is significantly different; therefore, the assumption of homogeneity has been violated. When sample sizes are large, small differences in group variances can produce a significant Levene's test.

The Levene's Test for Equality of Variances for 4th grade homogeneity of variance from Table 3 was found not to be violated for the writing scores $F(3, 279) = 1.47, p = .220$. Levene's Test for Equality of Variances was also found not to be violated for 7th grade $F(3, 1080) = .813, p = .487$. The results of the Levene's Test for Equality of Variances indicated that the variance between the test scores was not significant.

ANCOVA Analysis For 4th and 7th GradeWriting

The Analysis of Covariance (ANCOVA) shown in Table 4 was performed to evaluate whether 4th grade and 7th grade writing scores were equal across levels of a categorical independent variable (certification type) while controlling for effect of the continuous variable, the covariate, prior achievement. This analysis determined if significant differences between two groups existed by reducing the within-group variance. ANCOVA was also selected to correct for initial group differences (prior to group assignment that existed on the dependent variable among groups). In this study, participants could not be made equal by random assignment so the covariate was used to adjust scores.

Table 4

ANCOVA Test of Between -Subjects Effects of Writing 4th and 7th Grades

Grade Level 2013	Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
4	Corrected Model	37523983.829[a]	4	9380995.96	72.27	.000	.284
	Intercept	27460186.636	1	27460186.64	211.56	.000	.225
	W_SSC_12	37104564.601	1	37104564.60	285.86	.000	.282
	Certification Type	320648.108	3	106882.70	.823	.481	.003
	Error	94493741.579	728	129799.10			
	Total	8789487396.000	733				
	Corrected Total	132017725.408	732				
7	Corrected Model	43345641.660[b]	4	10836410.42	38.04	.000	.124
	Intercept	99116748.122	1	99116748.12	347.95	.000	.244

Table Continued

W_SSC_12	43337309.426	1	43337309.43	152.14	.000	.124
Certification Type	75472.763	3	25157.59	.088	.966	.000
Error	307363735.560	1079	284859.81			
Total	12127614029.000	1084				
Corrected Total	350709377.220	1083				

*p<0.05

ANCOVA Results for 4th and 7th Grade Writing

The Analysis of Covariance test was conducted to determine whether a statistically significant difference existed between the writing scores of students when controlling for prior achievement on the 2011 STAAR writing tests based on teachers' certification routes. Table 4 indicates there was not a significant difference in the effect of certification pathway on students' performance on grade 4 writing test when controlling for prior achievement F (3, 728) =.823, p =.481, $\acute{\eta}\rho^2$ =.003. The effect size was small essentially non-existent. The student scores on the 4th grade STAAR writing tests of traditionally-certified and alternatively-certified teachers were not different when controlling for prior achievement.

Also shown in Table 4 are the writing results for 7th grade. The analysis indicated that there was not a significant difference in the effect of certification pathway on students' performance on the grade 7 writing test when controlling for prior achievement F (3, 1079) =.088, p = .966, $\acute{\eta}\rho^2$ = .000. The student scores on the 7th grade

STAAR writing tests of traditionally-certified and alternatively-certified teachers were not different when controlling for prior achievement.

Descriptive Statistics of Alternative and Traditional Certification Programs

As shown in Table 5, the number of certifications for the Jamison Bill, which is an alternative certification program, is low. There was only 1 Jamison Bill participant in 5th grade, 66 participants in 6th grade, and 87 participants in 8th grade. In third grade, there were 456 alternative certified teachers, 91 out-of state teachers, and 865 traditionally-certified teachers. In 4th grade, there were 16,085 alternatively-certified teachers, 4,394 out-of-state teachers, 36,256 traditionally-certified teachers. In 5th grade, there were 22,799 alternatively-certified teachers, 5,148 out-of-state teachers, and 40,154 traditionally-certified teachers. In 6th grade, there were 26,319 alternatively-certified teachers, 8,128 out-of-state teachers, and 53,645 traditionally-certified teachers. In 7th grade, there were 47,242 alternatively-certified teachers, 10,263 out-of-state teachers, and 72,734 traditionally-certified teachers. In 8th grade, there were 55,226 alternatively-certified teachers, 8304 out-of-state teachers, and 65,266 traditionally-certified teachers. The Jamison Bill program is no longer in existence.

Table 5

Reading Between -Subjects Factors Certification

Grade Level 2013		Value Label	N
3	Certification Type	1 Alternative Program	390

		2	Alternative Program (Post-Bac)	56
		4	Out-of-State	91
		5	Traditional Program	865
		1	Alternative Program	13031
		2	Alternative Program (Post-Bac)	3054
4	Certification Type	4	Out-of-State	4394
		5	Traditional Program	36256
		1	Alternative Program	18967
		2	Alternative Program (Post-Bac)	3832
5	Certification Type	4	Out-of-State	5148
		5	Traditional Program	40154
		3	Jamison Bill	1
6	Certification Type	1	Alternative Program	21318

Table Continued

		2	Alternative Program (Post-Bac)	5001
		4	Out-of-State	8128
		5	Traditional Program	53645
		3	Jamison Bill	66
		1	Alternative Program	38202
7	Certification Type	2	Alternative Program (Post-Bac)	9040
		4	Out-of-State	10263
		5	Traditional Program	72734
		1	Alternative Program	43679
8	Certification Type	2	Alternative Program (Post-Bac)	11547
		4	Out-of-State	8304
		5	Traditional Program	65266
		3	Jamison Bill	67

Table 6

Descriptive Statistics of Reading Scaled Scores 2013 for Each Certification Program

Grade Level 2013	Certification Type	Mean	SD	N
3	Alternative Program	1400.37	312.19	390
	Alternative Program (Post-Bac)	1384.77	281.45	56
	Out–of-State	1389.32	166.36	91
	Traditional Program	1404.03	342.91	865
	Total	1401.29	323.25	1402
4	Alternative Program	1543.94	318.12	13031
	Alternative Program (Post-Bac)	1551.64	312.94	3054
	Out-of-State	1578.10	306.17	4394
	Traditional Program	1561.17	304.09	36256
	Total	1558.01	308.13	56735
5	Alternative Program	1588.86	312.85	18967
	Alternative Program (Post-Bac)	1598.64	301.29	3832
	Out-of-State	1619.44	304.38	5148
	Traditional Program	1608.70	311.72	40154
	Jamison Bill	1416.00	.	1
	Total	1603.42	311.06	68102
6	Alternative Program	1633.41	301.68	21318

Table Continued

	Alternative Program (Post-Bac)	1635.61	285.22	5001
	Out-of-State	1651.94	290.94	8128
	Traditional Program	1642.33	291.49	53645
	Jamison Bill	1672.35	378.10	66
	Total	1640.70	293.70	88158
7	Alternative Program	1675.23	289.04	38202
	Alternative Program (Post-Bac)	1668.57	258.00	9040
	Out-of-State	1674.75	261.21	10263
	Traditional Program	1673.95	252.86	72734
	Total	1674.02	264.97	130239
8	Alternative Program	1696.31	284.38	43679
	Alternative Program (Post-Bac)	1716.01	251.42	11547
	Out-of-State	1723.85	247.78	8304
	Traditional Program	1712.62	262.41	65266
	Jamison Bill	1684.45	85.09	67
	Total	1708.10	268.31	128863

ANCOVA Analysis For 3rd-8th Grade Reading

Table 7

ANCOVA Reading Scaled Score 2013

Grade Level 2013	Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
	Corrected Model	24343496.988[a]	4	6085874.25	69.66	.000	.166
	Intercept	12830625.715	1	12830625.72	146.87	.000	.095
	R_SSC_12	24308325.347	1	24308325.35	278.25	.000	.166
3	Certification_Type	117364.501	3	39121.50	.448	.719	.001
	Error	122044334.696	1397	87361.73			
	Total	2899380971.000	1402				
	Corrected Total	146387831.684	1401				
	Corrected Model	2465093064.413[b]	4	616273266.10	11966.37	.000	.458
	Intercept	508017872.323	1	508017872.32	9864.34	.000	.148
	R_SSC_12	2460254479.544	1	2460254479.544	47771.51	.000	.457
4	Certification_Type	469179.500	3	156393.17	3.04	.028	.000
	Error	2921620517.492	56730	51500.45			
	Total	143104515729.000	56735				
	Corrected Total	5386713581.906	56734				
	Corrected Model	3421344603.957[c]	5	684268920.79	14709.34	.000	.519
	Intercept	3554206.884	1	3554206.88	76.40	.000	.001
	R_SSC_12	3414759733.538	1	3414759733.538	73405.16	.000	.519
5	Certification Type	314468.138	4	78617.04	1.69	.149	.000
	Error	3167781216.893	68096	46519.34			
	Total	181675860524.000	68102				
	Corrected Total	6589125820.850	68101				

Table Continued

	Corrected Model	3713054799.075[d]	5	742610959.82	16822.55	.000	.488
	Intercept	296194912.371	1	296194912.37	6709.78	.000	.071
	R_SSC_12	3710555760.744	1	3710555760.744	84056.11	.000	.488
6	Certification Type	733878.003	4	183469.50	4.156	.002	.000
	Error	3891363721.394	88152	44143.79			
	Total	244917352507.000	88158				
	Corrected Total	7604418520.469	88157				
	Corrected Model	4732749085.884[e]	4	1183187271.471	34931.76	.000	.518
	Intercept	986189006.284	1	986189006.28	29115.69	.000	.183
	R_SSC_12	4732419028.957	1	4732419028.957	139717.28	.000	.518
7	Certification Type	85482.236	3	28494.08	.841	.471	.000
	Error	4411207112.564	130234	33871.39			
	Total	374116473727.000	130239				
	Corrected Total	9143956198.449	130238				
	Corrected Model	4456609937.467[f]	5	891321987.49	23827.86	.000	.480
	Intercept	265936290.009	1	265936290.01	7109.32	.000	.052
	R_SSC_12	4446387038.382	1	4446387038.382	118866.02	.000	.480
8	Certification Type	4819839.492	4	1204959.87	32.21	.000	.001
	Error	4820116975.115	128857	37406.71			
	Total	385249893538.000	128863				
	Corrected Total	9276726912.582	128862				

*p<.05

As shown in Table 6, The 3rd, grade reading mean and population size were (M = 1400.37, N = 390) for Alternatively-certified, for Alternative Program (Post-Baccalaureate) the mean and population size were (M =1384.77, N = 56), Out of State (M =1389.32, N =9 1), and traditionally-certified (M = 1404.03, N = 865) respectively. The differences in the number of teachers in alternatively-certified programs and traditionally-certified programs could be contributed to the fact that third grade is the first grade that students take the standardized test., STAAR test. More veteran teachers who are traditionally-certified may have been assigned 3rd grade.

ANCOVA Results for 3rd-8th Grade Reading

The Analysis of Covariance (ANCOVA) was selected in Table 7 to evaluate whether 3rd through 8th grade reading scores were equal across levels of a categorical independent variable (certification type) while controlling for effect of a continuous variable, a covariate, reading scores, 2012. The analysis was selected to determine if significant differences between two groups existed by reducing the within-group variance. Another use of ANCOVA was to correct for initial group differences (prior to group assignment that existed on the dependent variable among groups). In this study participants could not be made equal by random assignment, so the covariate prior achievement was used to adjust scores.

The effect size, $\eta\rho^2$, is the ratio of variance accounted for by an effect and that effect plus its associated error in an ANCOVA study. The norms for $\eta\rho^2$ are the following: small = .02, medium = .13, and large = .26.

As shown in Table 7, the 3rd grade reading ANCOVA test revealed that there was not a significant difference in effect of certification route on students' performance on the Grade 3 reading STAAR test when controlling for prior achievement $F(3, 1397) = .448$, $p = .719$, $\eta\rho^2 = .001$. The effect size was so small as to be essentially non-existent.

However, there was a significant difference in the effect of certification pathway on students' performance on the 4th grade STAAR reading test when controlling for prior achievement $F(3, 56{,}730) = 3.04$, $p = .028$, $\eta\rho^2 = .000$. The mean scaled scores of the traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 4th grade STAAR test. In 5th grade reading there was not a significant difference in the effect of certification route on students' performance on the Grade 5 reading STAAR test when controlling for prior achievement $F(4, 68{,}096) = 1.69$, $p = .149$, $\eta\rho^2 = .000$.

The 6th grade reading STAAR test indicated that there was a significant difference in the effect of certification pathway on students' performance on the 6th grade STAAR test when controlling for prior achievement $F(4, 88{,}152) = 4.16$, $p = .002$, $\eta\rho^2 = .000$. The mean scaled scores of the traditionally-certified teacher were significantly, but only slightly higher than the alternatively-certified teachers' student scores of the 6th grade STAAR test. In 7th grade reading there was not a significant difference in the effect of certification route on students' performance when controlling for prior achievement $F(3, 130{,}234) = .841$, $p = .471$, $\eta\rho^2 = .000$. As shown in Table 7, there was a significant difference in the effect of certification pathway on students' performance on the 8th grade STAAR reading test when controlling for prior achievement $F(4, 128{,}857) = 32.21$, $p = .000$, $\eta\rho^2 = .001$. The mean scaled scores of the traditionally-certified teachers were

significantly, but only slightly higher than the alternatively-certified student scores on the 8th grade STAAR test. The effect size, however, was so small as to be non-existent.

Table 8

Levene's Test for Equality of Variances for Reading Scaled Score 2013

Grade Level	F	df1	df2	Sig.
3	4.96	3	57425	.002
4	.222	3	68455	.881
5	1.39	4	87916	.236
6	27.08	3	130338	.000
7	29.82	4	128628	.000
8	.540	4	710	.707

*p<.05

Levene's Test for Equality of Variances for 3rd -8th Grade Reading Results

The Levene's Test for Equality of Variances was used to test the assumption of homogeneity of variance. The assumption was that there was no variance in the group means. When performing the analysis, the level of significance was set at $\alpha = .05$.

The Levene's Test for Equality of Variances were found not to be violated for the following grades: Grade 4 $F(3, 68455) = .222$, $p = .881$, Grade 5 $F(4, 87916) = 1.39$, $p =$

.236, Grade 8 F (4, 710) = .540, p = .707. This means population mean in the grade levels was the same because the level of significance was set at $\alpha = .05$ and the $p>.05$. Results for Grades 3, 6, and 7 indicated that variability was different for the reading scaled scores. Grade 3 F (3, 51425) = 4.96, p = .002, Grade 6 F (3, 130338), p = .000, and Grade 7 F (4, 128628) = .540, p = .000 respectively. The statistics associated with Grade 3, Grade 6, and Grade 7 indicated that differences in the variances of the populations existed as shown in Table 8 because homogeneity of variance was violated because the $p < .05$.

Descriptive Statistics for Alternative and Traditional Certified Teachers Programs

Table 9

Between-Subjects Factors of Certification Type For Math 2013

Grade Level			Value Label	N
3	Certification Type	1	Alternative Program	3
		2	Alternative Program (Post-Bac)	1
		4	Out-of-State	1
		5	Traditional Program	7
4	Certification Type	1	Alternative Program	19150
		2	Alternative Program (Post-Bac)	3839
		4	Out-of-State	5153
		5	Traditional Program	40312
5	Certification Type	1	Alternative Program	21185

Table Continued

		2	Alternative Program (Post-Bac)	4942
		4	Out-of-State	8119
		5	Traditional Program	53605
		3	Jamison Bill	67
6	Certification Type	1	Alternative Program	38194
		2	Alternative Program (Post-Bac)	9045
		4	Out-of-State	10270
		5	Traditional Program	72833
7	Certification Type	1	Alternative Program	43632
		2	Alternative Program (Post-Bac)	11523
		4	Out-of-State	8289
		5	Traditional Program	65123
		3	Jamison Bill	66
8	Certification Type	1	Alternative Program	276
		2	Alternative Program (Post-Bac)	97
		4	Out-of-State	42
		5	Traditional Program	299
		3	Jamison Bill	1

As shown in Table 9, Alternative Programs (Post-Baccalaureate) throughout Grades 3-8 mathematics had low participants in comparison to the other certification types. This may be attributed to a shortage of teachers with the academic credentials needed to obtain their license in the state of Texas. The number of participants under the

Jamison Bill was low in comparison to the various certification types. This is possibly due to the fact that the program was in the process of being phased out. The Jamison Bill program is no longer in existence.

Table 10

Descriptive Statistics Math Scaled Scores 2013

Grade Level	Certification Type	*M*	*SD*	*N*
3	Alternative Program	1959.67	741.19	3
	Alternative Program (Post-Bac)	1557.00	.	1
	Out-of-State	1503.00	.	1
	Traditional Program	2108.86	463.72	7
	Total	1975.08	514.57	12
4	Alternative Program	1620.81	313.77	19150
	Alternative Program (Post-Bac)	1626.92	293.61	3839
	Out-of-State	1648.82	313.63	5153
	Traditional Program	1638.99	308.56	40312
	Total	1633.97	309.72	68454
5	Alternative Program	1644.40	312.60	21185
	Alternative Program (Post-Bac)	1643.00	312.92	4942
	Out-of-State	1664.71	304.80	8119
	Traditional Program	1650.64	300.51	53605
	Jamison Bill	1731.16	442.48	67
	Total	1650.07	304.75	87918
6	Alternative Program	1611.08	348.83	38194
	Alternative Program (Post-Bac)	1621.38	317.96	9045
	Out-of-State	1595.97	336.89	10270
	Traditional Program	1599.03	326.21	72833
	Total	1603.87	333.36	130342
7	Alternative Program	1551.00	377.07	43632
	Alternative Program (Post-Bac)	1529.77	369.51	11523
	Out-of-State	1504.70	385.92	8289

Table Continued

	Traditional Program	1532.72	372.77	65123
	Jamison Bill	1480.80	263.63	66
	Total	1536.82	374.95	128633
	Alternative Program	1622.65	326.50	276
	Alternative Program (Post-Bac)	1627.09	304.05	97
8	Out-of-State	1598.14	298.52	42
	Traditional Program	1608.77	325.10	299
	Jamison Bill	1641.00	.	1
	Total	1616.03	320.61	715

Table 10 indicated in 3rd grade math Out-of-State certification participant was ($N = 1$) and the Jamison Bill for 5th grade, 7th grade, and 8th grade were low, ($N = 67$, $N = 66$, $N = 1$) respectively. As mentioned previously, the Jamison Bill certification program is no longer in existence, which may have resulted in the low number of participants who were certified by this program.

ANCOVA Analysis for 3rd-8th Grade Math

Table 11

ANCOVA Tests of Between-Subjects Effects of Math Scaled Scores

Grade Level	Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
3	Corrected Model	2550594.952[a]	4	637648.74	12.33	.003	.876
	Intercept	86313.990	1	86313.10	1.67	.237	.193
	M_SSC_12	2026957.559	1	2026957.56	39.20	.000	.848
	Certification Type	82823.302	3	27607.77	.534	.674	.186
	Error	361957.965	7	51708.28			
	Total	49724003.000	12				
	Corrected Total	2912552.917	11				
4	Corrected Model	2943301408.762[b]	4	735825352.19	13901.30	.000	.448
	Intercept	467813261.247	1	467813261.25	8837.99	.000	.114
	M_SSC_12	2937640606.757	1	2937640606.757	55498.27	.000	.448
	Certification Type	1084208.120	3	361402.71	6.83	.000	.000
	Error	3623149981.564	68449	52932.11			
	Total	189328132503.000	68454				
	Corrected Total	6566451390.325	68453				
5	Corrected Model	3295143913.747[c]	5	659028782.75	11897.17	.000	.404
	Intercept	332163534.597	1	332163534.60	5996.41	.000	.064
	M_SSC_12	3292017158.629	1	3292017158.629	59429.42	.000	.403
	Certification Type	1796926.407	4	449231.60	8.11	.000	.000
	Error	4869773672.891	87912	55393.73			
	Total	247540605061.000	87918				
	Corrected Total	8164917586.638	87917				
6	Corrected Model	3962005935.894[d]	4	990501483.97	12268.82	.000	.274

Table Continued

	Intercept	1505309733.849	1	1505309733.849	18645.47	.000	.125
	M_SSC_12	3954897437.086	1	3954897437.086	48987.22	.000	.273
	Certification Type	7147406.813	3	2382468.94	29.51	.000	.001
	Error	10522530418.940	130337	80733.266			
	Total	349776245471.000	130342				
	Corrected Total	14484536354.835	130341				
	Corrected Model	3451489438.719[e]	5	690297887.74	6068.19	.000	.191
	Intercept	579848408.708	1	579848408.71	5097.26	.000	.038
	M_SSC_12	3432290883.096	1	3432290883.096	30172.19	.000	.190
7	Certification Type	20017032.943	4	5004258.24	43.99	.000	.001
	Error	14632193913.065	128627	113756.78			
	Total	321891892111.000	128633				
	Corrected Total	18083683351.784	128632				
	Corrected Model	18231994.613[f]	5	3646398.92	46.87	.000	.248
	Intercept	6999797.782	1	6999797.78	89.97	.000	.113
	M_SSC_12	18178223.239	1	18178223.24	233.65	.000	.248
8	Certification Type	122499.302	4	30624.83	.394	.813	.002
	Error	55161761.513	709	77802.20			
	Total	1940665597.000	715				
	Corrected Total	73393756.126	714				

*p<0.5

ANCOVA Results for 3rd-8th Grade Mathematics

The Analysis of Covariance (ANCOVA) was selected in Table 11 to evaluate whether 3rd through 8th grade math scores were equal across levels of a categorical independent variable (certification type) while controlling for effects of continuous variable, a covariate, and math scores for 2012. ANCOVA was performed to determine

if significant differences between two groups existed by reducing the within-group variance. Another reason ANCOVA was selected was to correct for initial group differences that existed prior to group assignment that existed on the dependent variable among groups. In this study, participants could not be made equal by random assignment so the covariate was used to adjust scores.

The effect size, partial eta squared, is defined as the ratio of variance accounted for by an effect and that effect plus its associated error within an ANCOVA study. The norms for partial eta squared are the following: small = .02, medium = .13, and large = .26. The effect size explains the amount of variance accounted for in the sample.

The ANCOVA of the math scaled scores in Table 11 indicated that there were no significant differences in the effect of certification routes on students' performance on the 3rd grade math STAAR test when controlling for prior achievement F (3, 7) =. 534, p = .674, $\acute{\eta}\rho^2$ = .186. The effect size was medium, greater than .13.

There was a significant difference in the effect of certification pathways on students' performance on the 4th grade STAAR math test when controlling for prior achievement F (3, 68,449) = 6.83, p = 000, $\acute{\eta}\rho^2$ = .000. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 4th grade STAAR test. There was a significant difference in the effect of certification routes on students' performance on the 5th grade STAAR math test when controlling for prior achievement F (4, 87,912) = 8.11, p = .000, $\acute{\eta}\rho^2$ = .000. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 5th grade STAAR math test. There was a significant difference regarding

the effect of certification pathways on students' performance on the 6th grade STAAR math test when controlling for prior achievement F (3, 130,337) = 29.51, p = .000, $\acute{\eta}\rho^2$ =.001. The mean scaled scores of the alternatively-certified teachers were significantly, but only slightly higher than the traditionally-certified teachers' student scores on the 6th grade STAAR test. The effect size, however, was so small as to be non-existent. There was a significant difference in the effect of certification pathways on students' performance on the 7th grade STAAR math test when controlling for prior achievement F (4, 128,627) = 43.99, p =.000, $\acute{\eta}\rho^2$ = .000. The mean scaled scores of the alternatively-certified teachers were significantly, but only slightly higher than the traditionally-certified teachers' student scores on the 7th grade STAAR test. There was not a significant difference in the effect of certification routes on the students' performance on the 8th grade math STAAR test when controlling for prior achievement F (4, 709) = .394, p =.813, $\acute{\eta}\rho^2$ =.002. The effect size was very small, essentially non-existent.

Table 12

Levene's Test for Equality of Variance for Math Scaled Scores 2013

Grade Level	F	df1	df2	Sig
3	1.14	3	8	.389
4	1.74	3	68450	.157
5	6.81	4	87913	.000
6	48.14	3	130338	.000
7	70.22	4	128628	.000

8	.416	4	710	.797

*p<.05

Levene's Test of Equality of Variances for 3rd-8th Grade Mathematics Results

The Levene's Tests for Equality of Variances was performed to test the assumption of homogeneity of variance. The assumption was that there were no variances in the group means. When performing the analysis the level of significance was set at $\alpha = .05$.

The Levene's Tests for Equality of Variances were found not to be violated for the following analyses: 3rd grade math, $F(3, 8) = 1.14$, $p = .389$, 4th grade math, $F(3, 68450) = 1.74$, $p = .157$, and 8th grade math, $F(4, 710) = .416$, $p = .797$. Variability of the math scaled scores were the same in the aforementioned grade levels. In all, 5th grade math, 6th grade math, and 7th grade math scores violated the assumptions of the homogeneity of variances. These results revealed there were variances in the group means since $p<.05$. Specifically, 5th grade math was $F(4, 87913) = 6.81$, $p = .000$, 6th grade math was $F(3, 130338) = 48.14$, $p = .000$, and 7th grade $F(4, 12, 8628) = 70.22$, $p = .000$ as shown in Table 12.

Results for Research Question # 2

Research Question #2 was the following:

1. Is there a measurable difference between student scores in classrooms taught by alternatively-certified teachers and the scores of students taught by traditionally-certified teachers in rural, suburban, and urban schools in grades 3-8.

ANOVA Descriptives

The one-way Analysis of Variance test was performed to determine whether there was a measurable difference among the writing, reading, and math scores of students in classrooms taught by alternatively-certified teachers and the scores of students in classrooms taught by traditionally- certified teachers. The ANOVA analyses were specifically for rural, suburban, and urban schools in grades 3-8. In this analysis, the independent variable was the different communities in which the schools were located as shown in Table 13. The dependent variables were the reading, math, and writing scores of the students. The ANOVA test determined whether or not the group means were equal; however, this did not provide which group means had statistical differences. The ANOVA analysis produced an *F*-ratio that identified differences in the group means. The *F*-ratio detected when the experimental manipulation had an effect.

Table 13

ANOVA Descriptives

		N	M	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
	Charters	13044	1644.28	278.40	2.44	1639.50	1649.06	811	3693
	Independent Town	29127	1640.69	299.70	1.76	1637.25	1644.13	729	3895
	Major Suburban	164171	1655.43	289.36	.714	1654.03	1656.83	729	4229
Reading	Major Urban	87568	1630.67	309.83	1.05	1628.62	1632.72	623	3891
Scaled Score	Non-Metro	7920		N	3.010	1676.22	1688.36	811	3571
2013	Fast-Growing								
	Non-Metro Stable	31849	1644.85	299.97	1.68	1641.56	1648.14	811	3891
	Other CC Suburban	64632	1652.26	283.73	1.12	1650.07	1654.45	736	4229

Table Continued

	Other Central City	68396	1641.74	286.37	1.010	1639.60	1643.89	811	3779
	Rural	13874	1667.77	331.58	2.82	1662.26	1673.29	811	3779
	Total	480581	1647.45	294.35	.425	1646.62	1648.28	623	4229
	Charters	13044	1546.20	346.04	3.03	1540.26	1552.14	577	4579
	Independent Town	29127	1617.73	312.72	1.83	1614.14	1621.32	577	4257
	Major Suburban	164171	1592.09	343.94	.849	1590.43	1593.75	577	4675
	Major Urban	87568	1576.80	343.78	1.16	1574.52	1579.08	577	4579
	Non-Metro Fast-Growing	7920	1624.88	340.21	3.82	1617.39	1632.38	605	4025
Math Scaled Score 2013	Non-Metro Stable	31849	1618.48	321.26	1.80	1614.95	1622.00	577	5237
	Other CC Suburban	64632	1613.91	327.49	1.29	1611.39	1616.43	577	4901
	Other Central City	68396	1596.24	318.58	1.22	1593.85	1598.63	577	4144
	Rural	13874	1646.97	350.22	2.97	1641.15	1652.80	577	5126
	Total	480581	1597.01	335.57	.484	1596.06	1597.96	577	5237
	Charters	5384	3712.41	516.83	7.04	3698.60	3726.22	645	6390
	Independent Town	11009	3643.84	520.51	4.96	3634.11	3653.56	645	6390
	Major Suburban	65256	3778.89	558.31	2.19	3774.61	3783.18	499	6439
	Major Urban	35138	3641.54	537.82	2.89	3635.91	3647.16	499	6439
	Non-Metro Fast-Growing	3053	3850.31	559.62	10.13	3830.46	3870.17	645	6439
Writing Scaled Score 2013	Non-Metro Stable	12299	3665.22	505.06	4.55	3656.29	3674.14	645	6390
	Other CC Suburban	25144	3729.64	537.02	3.39	3723.00	3736.27	645	6439
	Other Central City	25275	3696.92	540.95	3.40	3690.25	3703.59	499	6439
	Rural	5267	3702.79	504.40	6.95	3689.16	3716.41	645	6439
	Total	187825	3717.33	544.15	1.26	3714.87	3719.79	499	6439

As shown in Table 13, the One-Way ANOVA descriptives showed that the 2013 reading scaled score for Rural was highest (M = 1667.77, SD = 331.28), when

compared to other communities. Major Urban had the lowest scores (M = 1630.67, SD = 309.83). Rural (M = 1646.97, SD = 350.22) was the highest for the 2013 math scaled score and Charter was the lowest score (M = 1546.20, SD = 346.04). Non-Metro Fast-Growing was the highest for the 2013 writing scaled score (M = 3850.31, SD = 559.62) and Major Urban had the lowest score (M = 3641.54, SD = 537.82).

Table 14

Levene's Test for Equality of Variances

	Levene Statistic	df1	df2	Sig.
Reading Scaled Score 2013	39.83	8	480572	.000
Math Scaled Score 2013	105.29	8	480572	.000
Writing Scaled Score 2013	22.91	8	187816	.000

*p<.05

Levene's Test for Equality of Variances Results

One of the first steps in the one-way ANOVA was to test the assumption of homogeneity of variance. The assumption was that there were no variances in the groups' means. The Levene's Test for Equality of Variances was used to test the assumption of homogeneity of variance. When performing the analysis, the level of significance was set at $\alpha = .05$.

Levene's Test for Equality of Variances was significant as shown in Table 14, 2013 reading scaled score F (8, 480572) = 39.83, p<.000, 2013 math scaled score, F (8,

480572) = 105.29, p<.000, and 2013 writing scaled score, F (8, 187816) = 22.91, p<.000. The homogeneity of variance was violated and significant differences between the variances exist because p<.05. The 2013 writing scaled scores, the 2013 reading scaled scores, and 2013 math scaled scores of traditionally-certified teachers and alternatively-certified teachers are significant different based on the following different socioeconomic areas: rural, urban, and suburban.

ANOVA Results

The one-way ANOVA, shown in Table 15, was performed; however, due to the violation of the homogeneity of variance, a more robust test was conducted to adjust for groups with larger sample sizes and large variances. Therefore, the adjusted F-ratio from the Welch F-test was used to determine whether group differences were statistically significant and a post hoc test was needed.

Table 15

One-way ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
2013	Between Groups	55873028.99	8	6984128.62	80.71	.000
	Within Groups	41583318613.22	480572	86528.80		
	Total	41639191642.21	480580			
Math Scaled Score 2013	Between Groups	159881618.31	8	19985202.29	177.10	.000
	Within Groups	53957635478.79	480572	112277.94		
	Total	54117517097.10	480580			

Table Continued

Table Continued

Writing Scaled Score 2013	Between Groups	611597652.71	8	76449706.59	261.05	.000
	Within Groups	55003379473.69	187816	292857.79		
	Total	55614977126.39	187824			

*p<.05

Table 16

Robust Tests of Equality of Means

		Statistic[a]	df1	df2	Sig.
Reading Scaled Score 2013	Welch	78.06	8	75523.70	.000
Math Scaled Score 2013	Welch	174.22	8	75600.12	.000
Writing Scaled Score 2013	Welch	257.97	8	29484.01	.000

*p<.05

Robust Test of Equality of Means Results

Since the homogeneity of variance was violated, an adjusted *F*-statistic was used, as shown in Table 16. The Welch *F*-test was instead used for the F-statistic, which was more powerful and conservative. The Welch *F*-ratio results were statistically significant at p<.05; therefore, a post hoc test was required. The reading scaled score for 2013 showed that there was a statistically significant difference in the group means of reading scaled scores for 2013 among students taught by traditionally and alternatively-certified teachers in rural, urban, and suburban areas F (8, 75523.70) = 78.06, p <.000. The significant difference was the student performance on the 2013 STAAR reading tests

when comparing the student scores of traditionally-certified and alternatively-certified teachers in rural, urban, and suburban areas.

The math scaled score for 2013 showed that there was a statistically significant difference in the group means among the math scaled scores for students taught by traditionally-and alternatively-certified teachers in rural, urban, and suburban areas F (8, 75600.12) =174.22, p <.000. The significant difference was the student performance on the 2013 STAAR math tests when comparing the student scores of traditionally-certified and alternatively-certified teachers in rural, urban, and suburban areas.

The writing scaled scores for 2013 showed that there was a statistically significant difference in the group means among the writing scaled scores for students taught by traditionally and alternatively-certified teachers in rural, urban, and suburban areas F (8, 29484.012) = 257.97, p<.000. The significant difference was the student performance on the 2013 STAAR writing tests when comparing the student scores of traditionally-certified and alternatively-certified teachers in rural, urban, and suburban areas.

The 2013 reading scaled score for Rural was the highest (M =1 667.77, SD = 331.28), when compared to the other communities. Major Urban had the lowest score (M = 1630.67, SD = 309.83). Rural (M = 1646.97, SD = 350.22) was the highest for the 2013 math scaled score and Charter was the lowest score (M = 1546.20, SD = 346.04). Non-Metro Fast-Growing was the highest for the 2013 writing scaled score (M = 3850.31, SD = 559.62) and Major Urban had the lowest score (M = 3641.54, SD = 537.82).

As a result of significant F-ratios on the Welch F-test for reading scaled score for 2013, math scaled score for 2013, and writing scaled score for 2013, the Scheffe post hoc

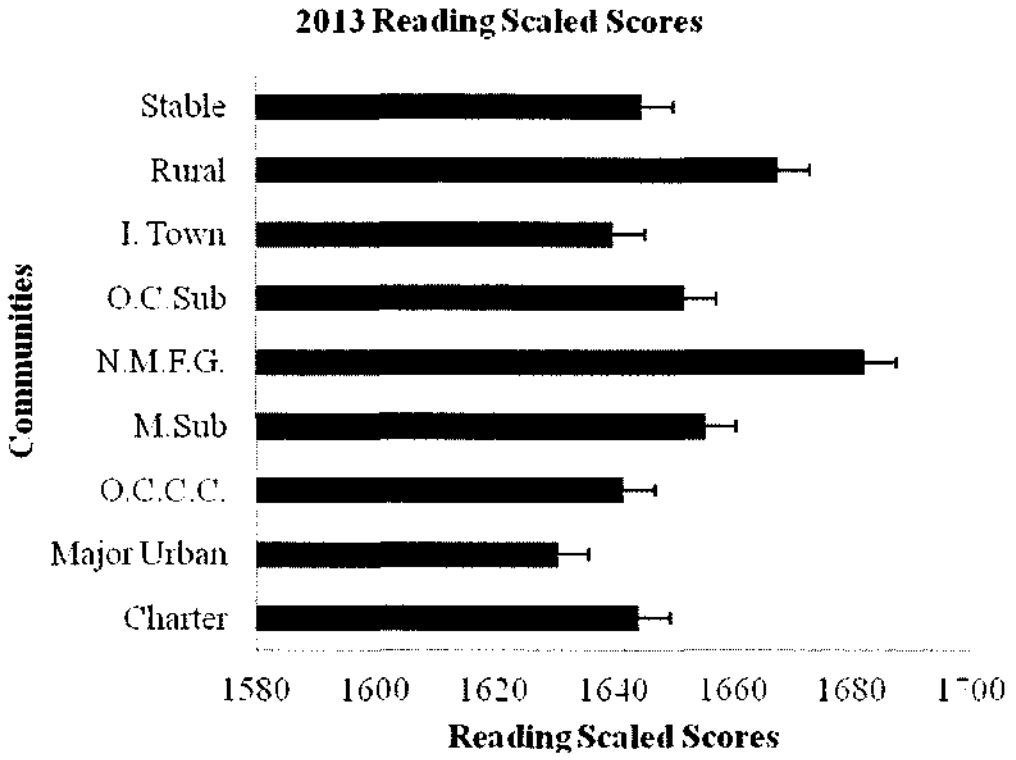
test was performed. This test was selected because it was conservative, and it corrected for alpha when used for simple and complex comparisons of students' performance on the reading, writing, and math 2013 STAAR tests.

Scheffe Post Hoc Test Results for 2013 Reading Scores

When performing the Scheffe post hoc test as shown in table 17, the results showed significant differences in means. Specifically, significant differences at $p < .05$ were revealed when comparing school community Charter, ($M = 1644.28$, $SD = 278.40$), which had a greater group mean in comparison to Major Urban and Other Central City school communities.

Figure 1

2013 Reading Scaled Scores



As shown in Figure 1, the 2013 student reading STAAR test scores of alternatively-certified and traditionally-certified were greater than the student scores of the alternatively and traditionally-certified teachers student scores in the following communities: these include Major Urban, (M = 1630.67, SD = 309.83) and Other Central City (M = 1641.74, SD = 286.37).

Major Suburban, (M = 1655.43, SD = 289.36), Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Non-Metro Stable (M = 1644.85, SD = 299.97) and Other C. C. Suburban (M = 1652.26, SD = 283.73) had greater group means in comparison to Charter which indicated the student scores on the 2013 reading STAAR test scores of traditionally-certified and alternatively-certified teachers were higher.

Independent Town (M = 1640, SD = 299.70) had a lower group mean in comparison to Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Other C.C. Suburban (M = 1652.26, SD = 283.73), Rural (M = 1667.77, SD = 331.58), and Charter (M = 1644.28, SD = 278.40). Major Suburban (M = 1655.43, SD = 289.36) had a lower group mean in comparison to Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Rural (M = 1667.77, SD = 331.58), and Independent Town (M = 1640.69, SD = 299.70). Major Suburban had a greater group mean in comparison to the following: Non-Metro Stable (M = 1644.85, SD = 299.97), Other Central City (M = 1641.74, SD = 286.37), Charters (M = 1644.28, SD = 278.40), and Independent Town (M = 1640, SD = 299.70).

Major Urban (M = 1630.67, SD =309.831) group mean was lower in comparison to the following: Major Suburban (M = 1655.43, SD = 289.36), Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Non-Metro Stable (M = 1644.85, SD = 299.97), Other C.C. Suburban (M = 1652.26, SD =283.73), Other Central City M = 1641.74, SD = 286.37), Rural (M = 1667.77, SD = 331.58), Charter (M = 1644.28, SD = 278.40), Independent Town (M = 1640.69, SD = 299.70) and Major Suburban, (M = 1655.43, SD = 289.36).

The group mean of Non-Metro Fast-Growing (M = 1682.29, SD = 275.62) was greater than that of the following communities: Major Urban (M = 1630.67, SD = 309.83), Non-Metro Stable (M = 1644.85, SD = 299.97), Other C.C. Suburban (M = 1652.26, SD = 283.73), Other Central Cities (M = 1641.74, SD = 286.37), Major Suburban (M = 1655.43, SD = 289.36).

Also, Major Urban (M = 1630.67, SD = 309.83) had a lower group mean in comparison to Non-Metro Fast-Growing which indicated alternatively and traditionally-

certified students scored lower on the reading 2013 STAAR test when compared to Non-Metro Fast-Growing alternatively and traditionally-certified teachers' student scores.

Non-Metro Stable (M = 1644.85, SD = 299.97) had a greater group mean in comparison to the following school community group means: Independent Town (M = 1617.73, SD = 312.72) and Major Urban (M = 1630.67, SD = 309.83).

Rural (M = 1667.77, SD = 331.58), Non-Metro Fast-Growing (M = 1682.29, SD=275.62) had a greater group mean in comparison to Non-Metro Stable.

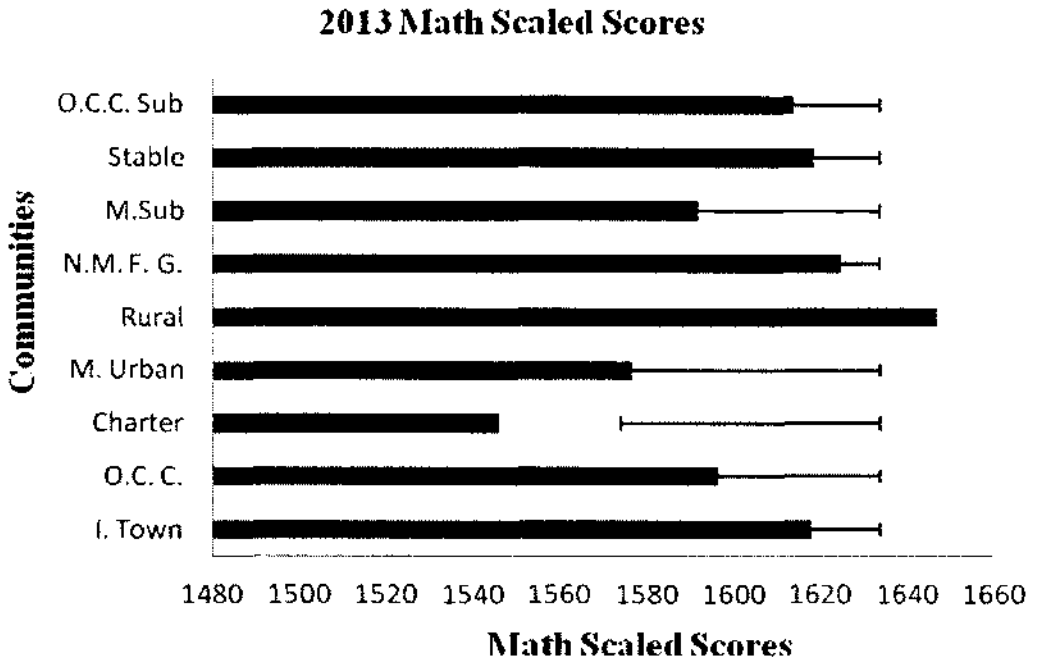
Other C.C. Suburban (M = 1652.26, SD = 283.73) had a greater group mean in comparison to Other Central City (M = 1641.74, SD = 286.37) and Major Urban (M = 1630.67, SD = 309.83). The following had a greater group mean than Other C.C. Suburban: Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Rural (M = 1667.77, SD = 331.58), and Major Suburban (M = 1655.43, SD = 289.36).

Scheffe Post Hoc Results for 2013 Math Scores

The 2013 math scaled score revealed that there was a statistically significant difference in the group means of the math scaled scores of students taught by traditionally-certified alternatively-certified teachers in rural, urban, and suburban areas. F (8, 75600.12) = 174.22, p <.000.

Figure 2

2013 Math Scaled Scores



As shown in Figure 2, the 2013 student math STAAR test scores of alternatively and traditionally-certified teachers of Independent Town (M = 1617.73, SD = 312.72) had a greater group mean in comparison to the following communities group means: Other Central City (M = 1596.24, SD = 318.58), Charters (M = 1546.20, SD = 346.04), and Major Urban (M = 1576.80, SD = 343.77). Also, the following communities that had a greater group mean in comparison to Independent Town's group mean: Rural (M = 1646.97, SD = 350.22) and Non-Metro Fast-Growing (M = 1624.88, SD = 340.21).

Major Suburban (M = 1592.09, SD = 343.94) had a lower group mean in comparison to the following communities' group means: Non-Metro Fast-Growing (M = 1624.88, SD = 340.21) and Non-Metro Stable (M = 1618.48, SD = 321.26).

Major Urban (M = 1576.80, SD = 343.78) had a lower group mean in comparison to the following communities' group means: Other C. C. Suburban (M =1613.91, SD =

327.49), Other Central City (M = 1596.24, SD = 318.58), and Rural (M = 1646.97, SD = 350.22).

Non-Metro Fast-Growing (M = 1624.88, SD = 340.21) had a greater group mean in comparison to the following communities' group means: Charter (M = 1546.20, SD = 346.04), Major Urban (M = 1576.80, SD = 343.78), Other Central City (M = 1596.24, SD = 318.58), and Independent Town (M = 1617.73, SD = 312.72). The group mean for Rural (M = 1646.97, SD = 350.18) was greater than that of Non-Metro Fast-Growing.

Other C. C. Suburban (M = 1613.91, SD = 327.49) had a greater group mean that of the following communities: Charters (M = 1546.20, SD = 346.04) and Other Central City (M = 1596.24, SD = 318.58). Other C.C. Suburban 2013 math STAAR test scores of alternatively-certified and traditionally-certified teachers were greater than the above student math scores of traditionally and alternatively-certified teachers in the above listed communities. Rural (M = 1646.97, SD = 350.18) had a greater group mean than Other C. C. Suburban which 2013 math STAAR test scores of traditionally and alternatively-certified teachers were lower than the Rural community's student math scores of traditionally-certified and alternatively-certified teachers.

Other Central City (M = 1596.24, SD = 318.58) had a greater group mean in comparison to the group means of the following: Rural (M = 1646.97, SD = 350.18), Non-Metro Stable (M = 1618.48, SD = 321.26), and Independent Town (M = 1617.73, SD = 312.72). The following communities had a lower group mean than Other Central City: Charter (M = 1546.20, SD = 346.04), Major Urban (M = 1576.80, SD = 343.78), and Major Suburban (M = 1592.09, SD = 343.94).

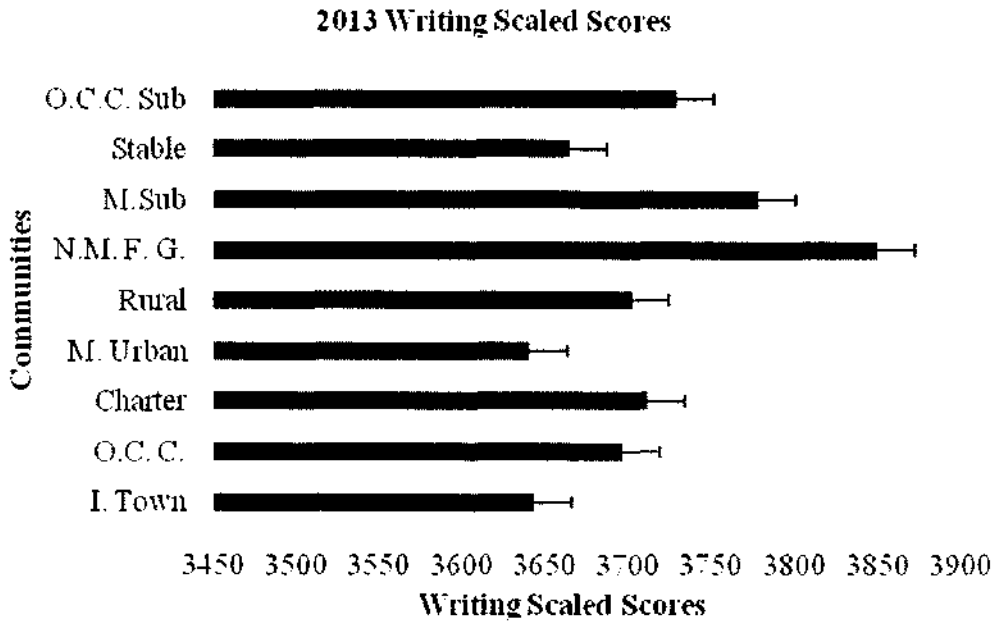
The group mean for Rural (M = 1646.97, SD = 350.22) was greater than that of the following communities: Major Urban (M = 1576.80, SD = 343.78), Non-Metro Fast-Growing (M = 1624.88, SD = 340.21), Non-Metro Stable (M = 1618.48, SD = 321.26), Other C.C. Suburban, (M = 1613.91, SD = 327.49), and Other Central City (M = 1596.24, SD = 318.59).

Scheffe Post Hoc Results for 2013 Writing Scores

The 2013 writing scaled score showed there was a statistically significant difference in the group means of the writing scaled scores of students taught by traditionally and alternatively- certified teachers in rural, urban, and suburban areas F (8, 29484.012) = 257.97, p<.000.

Figure 3

2013 Writing Scaled Scores



As shown in Figure 3, the 2013 student writing STAAR test scores of alternatively and traditionally-certified teachers of Charter (M = 3712.41, SD = 516.83) group mean was greater than the following communities' group means: Major Urban (M = 3641.54, SD = 537.82), Independent Town (M = 3643.84, SD = 520.51) and Non-Metro Stable (M = 3665.33, SD = 505.06). Non-Metro Fast-Growing (M = 3850.31, SD = 559.62) and Major Suburban (M = 3778.89, SD = 558.31) had a greater group mean in comparison to Charter.

Independent Towns (M = 3643.84, SD = 520.51) had a lower group mean in comparison to the group means of the following communities: Other C. C. Suburban (M = 3729.64, SD = 537.02), Other Central City (M = 3696.92, SD = 540.95), Rural (M =

3702.79, *SD* = 504.40), Major Urban (*M* = 3641.54, *SD* = 537.82) and Non-Metro Fast-Growing (*M* = 3850.31, *SD* = 559.62).

Major Suburban (*M* = 3778.89, *SD* = 558.32) had a greater group mean than those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06), Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M* = 3696.92, *SD* = 540.95), Rural (*M* = 3702.79, *SD* = 504.40), and Charters (*M* = 3712.41, *SD* = 516.83).

Major Urban (*M* = 3641.54, *SD* = 537.82) had a lower group mean than those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06) and Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02). Rural (*M* = 3702.79, *SD* = 504.40), Charters (*M* = 3712.41, *SD* = 516.83), Independent Town (*M* = 3643.84, *SD* = 520.57), and Major Suburban (*M* = 3778.89, *SD* = 558.31).

Non-Metro Fast-Growing (*M* = 3850.31, *SD* = 559.62) had a greater group mean in comparison to those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06), Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M* = 3696.92, *SD* = 540.95), Rural (*M* = 3702.79, *SD* = 504.40), and Charters (*M* = 3712.41, *SD* = 516.82).

Non-Metro Stable (*M* = 3665.22, *SD* = 505.06) had a lower group mean in comparison to those of the following communities: Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M* = 3696.92, *SD* = 540.95) and Rural (*M* = 3702.79, *SD* = 504.40).

Independent Town (*M* = 3643.84, *SD* = 520.57) and Major Urban (*M* = 3641.54, *SD* = 537.82) had group means lower than that of Non-Metro Stable (*M* = 3665.22, *SD* =505.06).

Other C.C. Suburban (M = 3729.64, SD = 537.02) had a group mean greater than those of the following communities: Non-Metro Stable (M = 3665.22, SD = 505.06), Other Central City (M = 3696.92, SD = 540.95), and Independent Town (M = 3778.89, SD = 558.31). Other Central City (M = 3696.92, SD = 540.95) had greater group mean in comparison to Major Urban (M = 3641.54, SD = 537.82).

Table 17

Scheffe' Post hoc Test

Dependent Variable	(I) Community Type	(J) Community Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Reading Scaled Score 2013	Charters	Independent Town	3.586	3.099	.995	-8.62	15.79
		Major Suburban	-11.157*	2.676	.026	-21.69	-.62
		Major Urban	13.609*	2.761	.002	2.74	24.48
		Non-Metro Fast-Growing	-38.016*	4.190	.000	-54.52	-21.51
		Non-Metro Stable	-.572	3.058	1.000	-12.61	11.47
		Other CC Suburban	-7.985	2.824	.434	-19.10	3.13
		Other Central City	2.533	2.810	.999	-8.53	13.60
		Rural	-23.497*	3.588	.000	-37.62	-9.37
	Independent Town	Charters	-3.586	3.099	.995	-15.79	8.62
		Major Suburban	-14.743*	1.870	.000	-22.11	-7.38
		Major Urban	10.023*	1.990	.001	2.19	17.86
		Non-Metro Fast-Growing	-41.602*	3.728	.000	-56.28	-26.92
		Non-Metro Stable	-4.158	2.385	.932	-13.55	5.23
		Other CC Suburban	-11.571*	2.076	.000	-19.75	-3.40
		Other Central City	-1.053	2.058	1.000	-9.16	7.05
		Rural	-27.083*	3.034	.000	-39.03	-15.13
	Major Suburban	Charters	11.157*	2.676	.026	.62	21.69
		Independent Town	14.743*	1.870	.000	7.38	22.11
		Major Urban	24.766*	1.231	.000	19.92	29.61
		Non-Metro Fast-Growing	-26.859*	3.384	.000	-40.19	-13.53
		Non-Metro Stable	10.585*	1.801	.000	3.49	17.68
		Other CC Suburban	3.172	1.366	.715	-2.21	8.55
		Other Central City	13.690*	1.339	.000	8.42	18.96
		Rural	-12.340*	2.601	.004	-22.58	-2.10
	Major Urban	Charters	-13.609*	2.761	.002	-24.48	-2.74
		Independent Town	-10.023*	1.990	.001	-17.86	-2.19
		Major Suburban	-24.766*	1.231	.000	-29.61	-19.92
		Non-Metro Fast-Growing	-51.624*	3.452	.000	-65.22	-38.03
		Non-Metro Stable	-14.181*	1.925	.000	-21.76	-6.60
		Other CC Suburban	-21.593*	1.525	.000	-27.60	-15.59
		Other Central City	-11.075*	1.501	.000	-16.99	-5.16

Table Continued

	Rural	-37.106*	2.688	.000	-47.69	-26.52
	Charters	38.016*	4.190	.000	21.51	54.52
	Independent Town	41.602*	3.728	.000	26.92	56.28
	Major Suburban	26.859*	3.384	.000	13.53	40.19
Non-Metro Fast-Growing	Major Urban	51.624*	3.452	.000	38.03	65.22
	Non-Metro Stable	37.443*	3.694	.000	22.90	51.99
	Other CC Suburban	30.031*	3.502	.000	16.24	43.82
	Other Central City	40.549*	3.491	.000	26.80	54.30
	Rural	14.519	4.143	.139	-1.80	30.83
	Charters	.572	3.058	1.000	-11.47	12.61
	Independent Town	4.158	2.385	.932	-5.23	13.55
	Major Suburban	-10.585*	1.801	.000	-17.68	-3.49
Non-Metro Stable	Major Urban	14.181*	1.925	.000	6.60	21.76
	Non-Metro Fast-Growing	-37.443*	3.694	.000	-51.99	-22.90
	Other CC Suburban	-7.412	2.014	.094	-15.34	.52
	Other Central City	3.106	1.995	.965	-4.75	10.96
	Rural	-22.925*	2.992	.000	-34.71	-11.14
	Charters	7.985	2.824	.434	-3.13	19.10
	Independent Town	11.571*	2.076	.000	3.40	19.75
	Major Suburban	-3.172	1.366	.715	-8.55	2.21
Other CC Suburban	Major Urban	21.593*	1.525	.000	15.59	27.60
	Non-Metro Fast-Growing	-30.031*	3.502	.000	-43.82	-16.24
	Non-Metro Stable	7.412	2.014	.094	-.52	15.34
	Other Central City	10.518*	1.614	.000	4.16	16.87
	Rural	-15.512*	2.752	.000	-26.35	-4.67
	Charters	-2.533	2.810	.999	-13.60	8.53
	Independent Town	1.053	2.058	1.000	-7.05	9.16
	Major Suburban	-13.690*	1.339	.000	-18.96	-8.42
Other Central City	Major Urban	11.075*	1.501	.000	5.16	16.99
	Non-Metro Fast-Growing	-40.549*	3.491	.000	-54.30	-26.80
	Non-Metro Stable	-3.106	1.995	.965	-10.96	4.75
	Other CC Suburban	-10.518*	1.614	.000	-16.87	-4.16
	Rural	-26.030*	2.739	.000	-36.82	-15.24
	Charters	23.497*	3.588	.000	9.37	37.62
	Independent Town	27.083*	3.034	.000	15.13	39.03
	Major Suburban	12.340*	2.601	.004	2.10	22.58
Rural	Major Urban	37.106*	2.688	.000	26.52	47.69
	Non-Metro Fast-Growing	-14.519	4.143	.139	-30.83	1.80
	Non-Metro Stable	22.925*	2.992	.000	11.14	34.71
	Other CC Suburban	15.512*	2.752	.000	4.67	26.35

Table Continued

		Other Central City	26.030*	2.739	.000	15.24	36.82
		Independent Town	-71.530*	3.530	.000	-85.43	-57.63
		Major Suburban	-45.889*	3.048	.000	-57.89	-33.89
		Major Urban	-30.602*	3.145	.000	-42.99	-18.22
		Non-Metro Fast-Growing	-78.684*	4.773	.000	-97.48	-59.89
	Charters	Non-Metro Stable	-72.277*	3.483	.000	-85.99	-58.56
		Other CC Suburban	-67.710*	3.216	.000	-80.38	-55.04
		Other Central City	-50.043*	3.201	.000	-62.65	-37.44
		Rural	-100.775*	4.087	.000	-116.87	-84.68
		Charters	71.530*	3.530	.000	57.63	85.43
		Major Suburban	25.641*	2.130	.000	17.25	34.03
		Major Urban	40.928*	2.266	.000	32.00	49.85
	Independent Town	Non-Metro Fast-Growing	-7.154	4.246	.944	-23.88	9.57
		Non-Metro Stable	-.747	2.717	1.000	-11.44	9.95
		Other CC Suburban	3.820	2.365	.956	-5.49	13.13
		Other Central City	21.487*	2.344	.000	12.26	30.72
		Rural	-29.245*	3.457	.000	-42.86	-15.63
		Charters	45.889*	3.048	.000	33.89	57.89
		Independent Town	-25.641*	2.130	.000	-34.03	-17.25
Math Scaled Score 2013		Major Urban	15.287*	1.402	.000	9.77	20.81
	Major Suburban	Non-Metro Fast-Growing	-32.794*	3.855	.000	-47.97	-17.61
		Non-Metro Stable	-26.387*	2.052	.000	-34.47	-18.31
		Other CC Suburban	-21.821*	1.556	.000	-27.95	-15.69
		Other Central City	-4.153	1.525	.492	-10.16	1.85
		Rural	-54.885*	2.963	.000	-66.55	-43.22
		Charters	30.602*	3.145	.000	18.22	42.99
		Independent Town	-40.928*	2.266	.000	-49.85	-32.00
		Major Suburban	-15.287*	1.402	.000	-20.81	-9.77
	Major Urban	Non-Metro Fast-Growing	-48.082*	3.932	.000	-63.56	-32.60
		Non-Metro Stable	-41.675*	2.193	.000	-50.31	-33.04
		Other CC Suburban	-37.108*	1.738	.000	-43.95	-30.27
		Other Central City	-19.441*	1.710	.000	-26.17	-12.71
		Rural	-70.173*	3.062	.000	-82.23	-58.12
		Charters	78.684*	4.773	.000	59.89	97.48
		Independent Town	7.154	4.246	.944	-9.57	23.88
	Non-Metro Fast-	Major Suburban	32.794*	3.855	.000	17.61	47.97
	Growing	Major Urban	48.082*	3.932	.000	32.60	63.56
		Non-Metro Stable	6.407	4.207	.970	-10.16	22.98
		Other CC Suburban	10.973	3.989	.477	-4.74	26.68

Table Continued

		Other Central City	28.641*	3.977	.000	12.98	44.30
		Rural	-22.091*	4.719	.005	-40.67	-3.51
		Charters	72.277*	3.483	.000	58.56	85.99
		Independent Town	.747	2.717	1.000	-9.95	11.44
		Major Suburban	26.387*	2.052	.000	18.31	34.47
		Major Urban	41.675*	2.193	.000	33.04	50.31
	Non-Metro Stable	Non-Metro Fast-Growing	-6.407	4.207	.970	-22.98	10.16
		Other CC Suburban	4.566	2.294	.861	-4.47	13.60
		Other Central City	22.234*	2.273	.000	13.28	31.19
		Rural	-28.498*	3.409	.000	-41.92	-15.08
		Charters	67.710*	3.216	.000	55.04	80.38
		Independent Town	-3.820	2.365	.956	-13.13	5.49
		Major Suburban	21.821*	1.556	.000	15.69	27.95
		Major Urban	37.108*	1.738	.000	30.27	43.95
	Other CC Suburban	Non-Metro Fast-Growing	-10.973	3.989	.477	-26.68	4.74
		Non-Metro Stable	-4.566	2.294	.861	-13.60	4.47
		Other Central City	17.668*	1.838	.000	10.43	24.91
		Rural	-33.064*	3.135	.000	-45.41	-20.72
		Charters	50.043*	3.201	.000	37.44	62.65
		Independent Town	-21.487*	2.344	.000	-30.72	-12.26
		Major Suburban	4.153	1.525	.492	-1.85	10.16
		Major Urban	19.441*	1.710	.000	12.71	26.17
	Other Central City	Non-Metro Fast-Growing	-28.641*	3.977	.000	-44.30	-12.98
		Non-Metro Stable	-22.234*	2.273	.000	-31.19	-13.28
		Other CC Suburban	-17.668*	1.838	.000	-24.91	-10.43
		Rural	-50.732*	3.120	.000	-63.02	-38.45
		Charters	100.775*	4.087	.000	84.68	116.87
		Independent Town	29.245*	3.457	.000	15.63	42.86
		Major Suburban	54.885*	2.963	.000	43.22	66.55
		Major Urban	70.173*	3.062	.000	58.12	82.23
	Rural	Non-Metro Fast-Growing	22.091*	4.719	.005	3.51	40.67
		Non-Metro Stable	28.498*	3.409	.000	15.08	41.92
		Other CC Suburban	33.064*	3.135	.000	20.72	45.41
		Other Central City	50.732*	3.120	.000	38.45	63.02
		Independent Town	68.575*	9.000	.000	33.13	104.02
		Major Suburban	-66.480*	7.673	.000	-96.70	-36.26
Writing Scaled Score 2013		Major Urban	70.874*	7.920	.000	39.68	102.06
	Charters	Non-Metro Fast-Growing	-137.902*	12.260	.000	-186.18	-89.62
		Non-Metro Stable	47.195*	8.843	.000	12.37	82.02

Table Continued

		Other CC Suburban	-17.226	8.127	.810	-49.23	14.78
		Other Central City	15.492	8.123	.888	-16.50	47.48
		Rural	9.625	10.488	.999	-31.68	50.93
		Charters	-68.575*	9.000	.000	-104.02	-33.13
		Major Suburban	-135.056*	5.576	.000	-157.01	-113.10
		Major Urban	2.298	5.911	1.000	-20.98	25.57
	Independent Town	Non-Metro Fast-Growing	-206.478*	11.069	.000	-250.07	-162.89
		Non-Metro Stable	-21.380	7.100	.337	-49.34	6.58
		Other CC Suburban	-85.801*	6.185	.000	-110.16	-61.45
		Other Central City	-53.084*	6.180	.000	-77.42	-28.75
		Rural	-58.951*	9.067	.000	-94.66	-23.25
		Charters	66.480*	7.673	.000	36.26	96.70
		Independent Town	135.056*	5.576	.000	113.10	157.01
		Major Urban	137.354*	3.581	.000	123.25	151.46
		Non-Metro Fast-Growing	-71.422*	10.021	.000	-110.88	-31.96
	Major Suburban	Non-Metro Stable	113.675*	5.320	.000	92.73	134.62
		Other CC Suburban	49.255*	4.017	.000	33.44	65.07
		Other Central City	81.972*	4.009	.000	66.18	97.76
		Rural	76.105*	7.752	.000	45.58	106.63
		Charters	-70.874*	7.920	.000	-102.06	-39.68
		Independent Town	-2.298	5.911	1.000	-25.57	20.98
		Major Suburban	-137.354*	3.581	.000	-151.46	-123.25
		Non-Metro Fast-Growing	-208.776*	10.211	.000	-248.99	-168.57
	Major Urban	Non-Metro Stable	-23.679*	5.670	.026	-46.01	-1.35
		Other CC Suburban	-88.099*	4.470	.000	-105.70	-70.50
		Other Central City	-55.382*	4.463	.000	-72.96	-37.81
		Rural	-61.249*	7.996	.000	-92.74	-29.76
		Charters	137.902*	12.260	.000	89.62	186.18
	Non-Metro Fast-	Independent Town	206.478*	11.069	.000	162.89	250.07
	Growing	Major Suburban	71.422*	10.021	.000	31.96	110.88
		Major Urban	208.776*	10.211	.000	168.57	248.99

Table Continued

	Non-Metro Stable	185.098*	10.942	.000	142.01	228.19
	Other CC Suburban	120.677*	10.372	.000	79.83	161.52
	Other Central City	153.394*	10.369	.000	112.56	194.23
	Rural	147.527*	12.310	.000	99.05	196.00
	Charters	-47.195*	8.843	.000	-82.02	-12.37
	Independent Town	21.380	7.100	.337	-6.58	49.34
	Major Suburban	-113.675*	5.320	.000	- 134.62	-92.73
	Major Urban	23.679*	5.670	.026	1.35	46.01
Non-Metro Stable	Non-Metro Fast-Growing	-185.098*	10.942	.000	- 228.19	- 142.01
	Other CC Suburban	-64.421*	5.955	.000	-87.87	-40.97
	Other Central City	-31.703*	5.950	.000	-55.13	-8.27
	Rural	-37.571*	8.911	.023	-72.66	-2.48
	Charters	17.226	8.127	.810	-14.78	49.23
	Independent Town	85.801*	6.185	.000	61.45	110.16
	Major Suburban	-49.255*	4.017	.000	-65.07	-33.44
	Major Urban	88.099*	4.470	.000	70.50	105.70
Other CC Suburban	Non-Metro Fast-Growing	-120.677*	10.372	.000	- 161.52	-79.83
	Non-Metro Stable	64.421*	5.955	.000	40.97	87.87
	Other Central City	32.717*	4.820	.000	13.74	51.70
	Rural	26.850	8.201	.218	-5.44	59.14
	Charters	-15.492	8.123	.888	-47.48	16.50
	Independent Town	53.084*	6.180	.000	28.75	77.42
	Major Suburban	-81.972*	4.009	.000	-97.76	-66.18
	Major Urban	55.382*	4.463	.000	37.81	72.96
Other Central City	Non-Metro Fast-Growing	-153.394*	10.369	.000	- 194.23	- 112.56
	Non-Metro Stable	31.703*	5.950	.000	8.27	55.13
	Other CC Suburban	-32.717*	4.820	.000	-51.70	-13.74
	Rural	-5.867	8.197	1.000	-38.15	26.41
	Charters	-9.625	10.488	.999	-50.93	31.68
	Independent Town	58.951*	9.067	.000	23.25	94.66
	Major Suburban	-76.105*	7.752	.000	- 106.63	-45.58
	Major Urban	61.249*	7.996	.000	29.76	92.74
Rural	Non-Metro Fast-Growing	-147.527*	12.310	.000	- 196.00	-99.05
	Non-Metro Stable	37.571*	8.911	.023	2.48	72.66

Table Continued

Other CC Suburban	-26.850	8.201	.218	-59.14	5.44
Other Central City	5.867	8.197	1.000	-26.41	38.15

*. The mean difference is significant at the p<.05 level.

SUMMARY OF RESULTS

The ANCOVA statistical test determined if there were any significant differences in students' scores according to certification types. ANCOVA evaluated whether the population means of a dependent variable (DV) was equal across levels of a categorical independent variable (IV), while controlling for the effect of a continuous variable, a covariate. Covariance is a measure of how much two variables change together and how strong a relationship exists between them. The covariate in this study was controlling for prior achievement, which was the students' performance on the 2012 math, 2012 reading, and 2012 writing STAAR tests. The independent variable was certification types and the dependent variable was the 2013 math, reading, and writing scaled scores for the STAAR tests. The effect size, $\acute{\eta}\rho^2$, is interpreted in the ANCOVA analysis as the ratio of variance accounted for by an effect and that effect plus any associated error within an ANCOVA study. The norms for partial eta squared, $\acute{\eta}\rho^2$, are the following: small=.02, medium = 13, and large =. 26.

The first step before performing the ANCOVA analysis was to conduct the Levene's Test of Equality of Variances. The assumption was that there were no variances in the group means. The Levene's Test for Equality of Variances was used to test the assumption of homogeneity of variance. When performing the analysis, the level of significance was set at α=.05. A significant result with a p<.05 indicates that the variances are significantly different; therefore, the assumption of homogeneity has been

violated. When sample sizes are large, small differences in group variances can produce a significant Levene's test.

Levene's Test for Equality of Variances Results

The Levene's Test for Equality of Variances was violated when testing for equality in variances of the group means for 5th and through 7th grade math scores. The test for homogeneity of variance was also violated for 3rd grade, 6th grade, and 7th grade reading scores. The test for homogeneity of variance was not violated for 4th and 7th grade writing.

ANCOVA Results for Writing

The ANCOVA results revealed that there were no statistically significant differences in the effect of certification route on students' performance on the grade 4 and grade 7 writing tests when controlling for prior achievement. The effect size was so small that there is no practical difference. The student scores of the 4th and 7th grade traditionally-certified and alternatively-certified teachers were not different when controlling for prior achievement.

ANCOVA Results for Reading

The ANCOVA results revealed that there were not statistically significant differences in the effect of certification pathways of 3rd, 5th and 7th grade reading when controlling for prior achievement. The effect sizes were so small as to be essentially non-existent. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers mean scores on the 4th, 6th and

8th grade STAAR reading tests. Again, the effect sizes were so small as to be essentially non-existent.

ANCOVA Results for Mathematics

The ANCOVA results revealed there were significant differences in 4th through 7th grade math when controlling for prior achievement. The mean scale scores of traditionally-certified teachers were significantly, but only slightly higher than the mean scale scores of alternatively-certified teachers on the 4th and 5th grade STAAR math tests. The mean scale scores of alternatively-certified teachers were significantly, but only slightly higher than the mean scale scores of traditionally-certified teachers on the 6th and 7th grade math test. The effect sizes were small as to be essentially non-existent.

The ANCOVA results revealed that there was no statistically significant difference in the effect of certification route on 8th grade students performance on the 2013 STAAR math test of traditionally-certified and alternatively –certified teachers when controlling for prior achievement. The effect size was so small as to be essentially non -existent.

ANOVA Analysis Results

The one-way ANOVA revealed that there were statistically significant differences in group means of traditionally-certified and alternatively certified students' scores in rural, urban, and suburban areas with a $p<.05$. The 2013 reading scaled score for Rural was the highest ($M = 1667.77$, $SD = 331.28$), when compared to the other communities. Major Urban had the lowest score ($M = 1630.67$, $SD = 309.83$). Rural ($M = 1646.97$, $SD = 350.22$) was the highest for the 2013 math scaled score and Charter was the lowest

score (M = 1546.20, SD = 346.04). Non-Metro Fast-Growing was the highest for the 2013 writing scaled score (M = 3850.31, SD = 559.62) and Major Urban had the lowest score (M = 3641.54, SD = 537.82).

The Welch F-test was performed to adjust for the means because the Levene's Homogeneity of Variance test was violated with a $p<.05$. Since the Welch's F-test was statistically significant for reading scaled scores, writing scaled scores, and math scaled scores, the Scheffe post hoc test was performed to determine what community had statistically significant differences in the group means with a $p< .05$. The Scheffe test revealed traditionally-certified and alternatively-certified teachers' students performed differently on the 2013 reading, writing, and math STAAR tests based on their socioeconomic demographics. The Welch's F-test did not identify specifically what the differences were; however, the test did reveal there were statistically significant differences in the 2013 reading, writing, and math scores of traditionally-certified and alternatively-certified teachers in the urban, rural, and suburban areas. A thorough discussion of the findings is presented in Chapter V.

CHAPTER V
DISCUSSION

The intent of this quantitative study is to present whether there are any significant differences in the test scores on end-of the-year summative assessments when comparing the student scores in classrooms taught by traditionally-certified teachers and those students taught by alternatively-certified teachers. This study aims to determine whether or not teacher licensure route has an effect on student scores. This chapter summarizes information from the previous chapter, provides a discussion of the findings, implications for educators, and recommendations for further research.

Chapter I described the history of teacher licensure and how the authority for licensing teachers passed from ecclesiastical to civil authorities (Angus, 2001; Feistritzer, 2008). As the licensure of teachers evolved during the 19th and early 20th centuries, the responsibility for teacher licensure and professional training was transferred from local governments to the states. During this time, completion of teacher education programs and coursework became necessary to obtain licensure for prospective teachers (Kaplan & Owings, 2011).

Soon alternative routes to teaching became another way for potential candidates to enter the field. Alternative teacher certification programs provided opportunities for school divisions to recruit and hire teachers who had subject matter knowledge, but who lacked the educational coursework required by traditional certification programs. Alternative certification programs offered prospective teachers the opportunity to complete coursework and gain experience as the teacher of record. Generally speaking, there are three common distinguishing features between traditional and alternative

certification programs: the controlling institution, the sequencing of certification requirements, and the terminology used to describe the teacher preparation programs (Evans, 2010). Traditional programs require experience obtained by completing student teaching or a practicum. Also, the teachers must pass the state examination before becoming the teacher of record.

High-stakes testing is the process of attaching significant consequences to standardized test performance, with the goal of increasing teacher effectiveness and student achievement (Berliner, Glass, Nichols, 2012; Haertel, Herman, 2005; Ryan, 2004). The rationale behind high-stakes testing is that teachers will work harder and students will be motivated to learn if incentives or penalties are attached to student performance on state standardized tests. President Bush signed the No Child Left Behind (NCLB) in January of 2002. It quickly changed the impact of federal influence on accountability for school divisions' student performance on assessments linked to state standards. Sanctions were enacted for poor performance, including schools being denied accreditation or being forced to monitor and improve the academic achievement of students in public schools. Race to the Top (RTTT) was authorized under the American Recovery and Reinvestment Act of 2009 during President's Obama first term. RTTT focused on student growth over time and included multiple measures for evaluating teachers. Like NCLB, RTTT emphasized improving teacher quality, thus closing student achievement gaps. To receive funds from this federal grant, states had to submit proposals including student growth as one of the measures encompassing the teacher evaluation system. Delaware and Tennessee were the first states to receive funding from this grant.

The current state standardized test for Texas is the State of Texas Assessment of Academic Readiness (STAAR), which was first administered to students in the spring of 2012. Students in grades 3 through 8 are tested in reading and mathematics. Writing is tested in grades 4 and 7 (Texas Education Agency, 2013).

There are five routes that prospective teachers can take to obtain their teacher license in Texas. The first route is a traditional university-based program that has been approved by the State Board of Educator Certification (SBEC). This pathway usually results in a baccalaureate degree. The second route for licensure is an alternative program approved by the SBEC. These programs usually involve the completion of coursework, mentoring, supervision, and professional development. A third route is certification by examination, in which the teacher obtains licensure by taking an exam in another content area and successfully passes the test. This route is for teachers who already have a Texas teaching license. The fourth route is recognition of an out an out-of-state certification. The final route for certification is emergency certification. This is a right given to school divisions to hire personnel in critical areas established by the SBEC (Texas Education Agency, 2013).

Chapter II compared traditional and alternative certification teacher program studies. From this examination of the literature it is clear that alternatively-certified teachers' students performed well on state certified tests when their teachers received professional development, mentoring, supervision, and teaching experience (Heilig & Jez, 2010).

New York researchers found that teacher certification from some traditional programs was a significant predictor of student performance (Board of Regents, 2008;

Darling-Hammond, 2010). Another study stated that student achievement was most enhanced by having a certified teacher who had graduated from a university pre-service program and completed more than two years of teaching experience (Boyd et al., 2007; Darling-Hammond, 2010).

Clotfelder, Ladd, and Vigdor (2007) conducted a study in North Carolina comparing teacher credentials and student achievement in high school. The study compared teachers with regular licensure who completed a state traditional program. The alternatively-certified teachers with some teaching experience appeared to be no less effective than teachers with a traditional license. This is attributed to the fact that many teachers received on-the-job training during their first two years of teaching. Similarly, this study found that alternatively-certified teachers performed not as well as traditionally-certified teachers in high school content areas.

Blackstone (2010) performed a quantitative correlation study of teacher preparation program effect on student achievement in rural school settings. The study compared test scores of traditionally-certified and alternatively-certified teachers' student scores on the seventh and eighth grades' end-of-course mathematics test. The researcher revealed no significant difference in measured effect on student achievement among traditionally and alternatively-trained teachers.

Burns, Gansle, and Noell (2012) found mixed results when performing a hierarchical linear modeling study of student achievement among traditionally-certified and alternatively-certified teachers in Louisiana schools. The alternative program, called the Masters Alternate Certification Program I, was the only alternative-teacher program that scored comparably to the scores of the traditional programs.

Nunnery, Kaplan, Owings, and Pribesh (2010) performed a study in Florida comparing teachers funded by Troops to Teachers with traditionally-certified teachers. The study compared student performance between the two teacher types on the 2003 and 2004 Florida Comprehensive Test (FCAT) in reading and mathematics. Teachers' years of experience were matched in this study to eliminate selection bias. The study revealed that the students of Troop to Teachers outperformed students of traditionally-certified teachers (Nunnery et al., 2010). This study confirmed that the students of alternatively-certified teachers can perform comparably or better than the students of traditionally-certified teachers on standardized tests.

The studies available on whether or not teacher certification routes affect student achievement are ultimately inconclusive. More studies are needed using longitudinal data that link teachers directly to student performance on state standardized tests. These studies should include teachers' years of experience, certification routes, and socioeconomic status of the schools. To determine a definitive answer on whether teachers' certification routes affect students' performance on high-stakes tests, state departments of education should collect data and link teacher licensure pathway to students' performance on state assessments. This study was an attempt to expand the research on teacher licensure.

Research Design

The researcher used an ex-post facto design in this quantitative study. This was the appropriate design because there is no direct manipulation of the independent variable (Leedy & Ormrod, 2005). This was a causal comparative study. This design was used because it attempted to determine cause for existing conditions or pre-existing differences

in groups. In this case, the cause and effect had already occurred. Ex-post facto was also the selected research design because it identified cause-effect relationships by involving two or more groups in the study. Individuals were already in each group before the study began. The sample was not randomly selected.

Ex-post facto design was also the most appropriate design for the study because the Texas Education Agency sampled from the population of alternatively-certified and traditionally-certified teachers who have taught in tested content areas in 3rd through 8th grades. The design was suitable because the direct manipulation and collection of the data were done by the Texas Education Agency's accountability department, which resulted in the entire universe of the population of alternatively- and traditionally-certified teachers who taught students in 3rd through 8th grades. The independent variable was not manipulated by the researcher, but by the accountability department of the Texas Education Agency. This ensured the validity of the collection of the data.

Research Questions

The research questions included the following:

1. Are there significant differences between alternatively-certified and traditionally-certified teachers' student scores on the following content area STAAR tests?
    - Grades 3-8 mathematics
    - Grades 3-8 reading
    - Grades 4 and 7 writing only
2. Is there a measurable difference between student scores in classrooms taught by alternatively-certified teachers and student scores from traditionally-certified teachers in rural, suburban, and urban schools in grades 3 through 8?

The statistical analyses used in this study were the Analysis of Covariance (ANCOVA) and One -Way Analysis of Variance (ANOVA). ANCOVA evaluated whether or not the population means of a dependent variable (DV) are equal across levels of a categorical independent variable (IV) while controlling for the effects of other continuous variable, a covariate. When performing ANCOVA, the DV means were adjusted to what they would be if all groups were equal on the covariate. Covariance is a measure of how much two variables change together and how strong of a relationship exists between them. Covariates can be a cofounding variable that can influence the DV. The covariate in this study controlled for prior achievement. ANCOVA was used in this study to determine whether or not significant differences between two groups exist by reducing the within-group variance. Another use of ANCOVA was to correct for initial group differences (prior to group assignment) that exist regarding DV among groups. In this situation, participants cannot be made equal through random assignment, so the covariate will be used to adjust scores (Field, 2011).

One-way Analysis of Variance (ANOVA) was used to determine if there were significant differences in the group means of traditionally-certified and alternatively-certified students' scores in rural, suburban, and urban areas. The independent variable IV was the communities in which the schools were located and the dependent variables DV were the math scaled scores, writing scaled scores, and reading scaled scores of the students.

The assumptions made when conducting the study were the following: the Texas Education Agency's database screened out data for those students who took the alternative test due to special needs; the students' scores from the two teacher types came

from the same schools; students who tested more than once on the same test data were removed; schools were designated as urban, rural, and suburban; and alternatively-certified teachers received the same opportunities for professional development and guidance from instructional leaders when teaching the curriculum and planning.

Summary of Results

The ANCOVA statistical test determined if there were any significant differences in students' scores according to certification types. ANCOVA evaluated whether the population means of a dependent variable DV was equal across levels of a categorical independent variable IV, while controlling for the effect of a continuous variable such as a covariate. The independent variables are the certification types which were manipulated by the Texas Education Agency when collecting the sample. The dependent variables are the students' scores on the 2013 math, 2013 reading, and 2013 writing STAAR tests. Covariance is a measure of how much two variables change together and how strong a relationship exists between them. The covariate in this study was controlling for prior achievement that was the students' performance on the 2012 math, reading, and writing STAAR tests. The effect size, $\acute{\eta}\rho^2$, is interpreted in the ANCOVA analysis as the ratio of variance accounted for by an effect and that effect plus any associated error within an ANCOVA study. The norms for partial eta squared, $\acute{\eta}\rho^2$, are the following: small = .02, medium = .13, and large = .26.

The first step before performing the ANCOVA analysis was to conduct the Levene's Test for Equality of Variances. The assumption was that there were no variances in the group means. The Levene's Test for Equality of Variances was used to test the assumption of homogeneity of variance. When performing the analysis, the level

of significance was set at α=.05. A significant result with a $p<.05$ indicates that the variances are significantly different; therefore, the assumption of homogeneity has been violated. When sample sizes are large, small differences in group variances can produce a significant Levene's test.

Levene's Test of Equality of Variances Results

The Levene's Test for Equality of Variances was violated when testing for equality in variances of the group means for 5th through 7th grade math scores. The test for homogeneity of variance was also violated for 3rd grade, 6th grade, and 7th grade reading scores. The Levene's Test for Equality of Variances was not violated when testing for equality in variances of the group means for 4th and 7th grade writing scores. This means variance in test scores was significantly different in the math and reading grade levels.

ANCOVA Results

The ANCOVA results revealed that there were no statistically significant differences in the effect of certification route on students' performance on the grade 4 and grade 7 writing tests when controlling for prior achievement.

The ANCOVA results revealed that there were not statistically significant differences in the effect of certification pathways on 3rd, 5th and 7th students' performance on the reading STAAR tests when controlling for prior achievement. The effect sizes for 3rd, 5th, and 7th grades were very small, essentially non-existent, because there was a very large sample size and the ANCOVA did not produce significant results. There was a significant difference in the effect of certification pathway on students'

performance on the 4th grade STAAR reading test when controlling for prior achievement F (3, 56,730) = 3.04, p = .028, ήp² =. 000. The mean scaled scores of the traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 4th grade STAAR test. There was a significant difference in the effect of certification pathway on students' performance on the 6th grade STAAR reading test when controlling for prior achievement F (4, 88,152) = 4.16, p = .002, ήp² = .000. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher, than the alternatively-certified teachers' student scores on the 6th grade STAAR test. There was a significant difference in the effect of certification pathway on students' performance on the 8th grade STAAR reading tests when controlling for prior achievement F (4, 128,857) = 32.21, p = .000, ήp² =. 001. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 8th grade STAAR test. The effect size for 8th grade STAAR reading test was so small, essentially non-existent.

The ANCOVA results revealed there were significant differences in 4th through 7th grade math when controlling for prior achievement. There was significant difference in the effect of certification pathway on 4th grade students' performance on the 4th grade STAAR math test when controlling for prior achievement F (3, 68, 449) = 6.83, p =. 000, ήp² = .000. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 4th grade STAAR math test. There was significant difference in the effect of certification pathway on 5th grade students' performance on the 5th grade STAAR math test when

controlling for prior achievement F (4, 87,912) = 8.11, p = .000, ηp^2 = .000. The mean scaled scores of traditionally-certified teachers were significantly, but only slightly higher than the alternatively-certified teachers' student scores on the 5th grade STAAR math test. There was a significant difference in the effect of certification pathway on 6th grade students' performance on the STAAR math test when controlling for prior achievement F (3, 130, 337) = 29.51, p = .000, ηp^2 =. 001. The mean scaled scores of alternatively-certified teachers were significantly, but only slightly higher than the traditionally-certified teachers' scores on the 6th grade STAAR test, however, the effect size was essentially non-existent. There was a significant difference in the effect of certification pathway on students' performance on the 7th grade STAAR math test when controlling for prior achievement F (4,128, 627) = 43.99, p = .000, ηp^2 = .000. The mean scaled scores of alternatively-certified teachers were significantly, but only slightly higher than traditionally-certified teachers' student scores on the 7th grade STAAR test, however, the effect size was essentially non-existent.

The ANCOVA results revealed that there were no significant differences in the effect of certification pathway on 3rd and 8th grade students' performance on the STAAR math tests when controlling for prior achievement. The effect size, for 3rd grade math was medium, $\eta \rho^2$ = .186. The effect size for 8th grade math was ηp^2 = .002, very small, essentially non-existent.

Robust Test of Equality of Means Results

The one-way ANOVA revealed that there were statistically significant differences in group means of traditionally-certified and alternatively- certified students' scores in rural, urban, and suburban areas with a $p<.05$. As shown in Table 13, the One-Way

ANOVA descriptives showed that all the 2013 reading scaled scores for Rural were the highest (M = 1667.77, SD = 331.28), when compared to the other communities. Major Urban had the lowest scores (M = 1630.67, SD = 309.83). Rural (M = 1646.97, SD = 350.22) was the highest for the 2013 all the math scaled scores and Charter was the lowest score (M = 1546.20, SD = 346.04). Non-Metro Fast-Growing was the highest for the all the 2013 writing scaled scores (M = 3850.31, SD = 559.62) and Major Urban had the lowest score (M = 3641.54, SD = 537.82).

The Welch F-test was performed to adjust for the means because the Levene's Test for Equality of Variances was violated with a $p<.05$. This means that there was variability in the student test scores. The Welch F-test was conducted to adjust for groups with larger sample sizes and large variances. The Welch F-test was instead used for the F-statistic, which was more powerful and conservative. The Welch F-ratio results were statistically significant at $p<.05$. The reading scaled score for 2013 showed that there was a statistically significant difference in the group means of reading scaled scores for 2013 among students taught by traditionally and alternatively-certified teachers in rural, urban, and suburban areas F (8, 75523.70) = 78.06, $p <.000$. The significant difference was the student performance on the 2013 STAAR reading tests when comparing the student scores of traditionally-certified and alternatively-certified teachers in rural, urban, and suburban areas. The 2013 reading scaled score for Rural was the highest (M = 1667.77, SD = 331.28), when compared to the other communities. Major Urban had the lowest score (M = 1630.67, SD = 309.83).

The math scaled score for 2013 showed that there was a statistically significant difference in the group means among the math scaled scores for students taught by

traditionally-and alternatively-certified teachers in rural, urban, and suburban areas F (8, 75600.12) 174.22 = p <.000. The significant difference was the student performance on the 2013 STAAR math tests when comparing the student scores of traditionally-certified and alternatively-certified teachers in rural, urban, and suburban areas. Rural (M = 1646.97, SD = 350.22) was the highest for the 2013 math scaled score and Charter was the lowest score (M = 1546.20, SD = 346.04).

The writing scaled scores for 2013 showed that there was a statistically significant difference in the group means among the writing scaled scores for students taught by traditionally and alternatively-certified teachers in rural, urban, and suburban areas F (8, 29484.012) = 257.97, p<.000. The significant difference was the student performance on the 2013 STAAR writing tests when comparing the student scores of traditionally-certified and alternatively-certified teachers in rural, urban, and suburban areas. Non-Metro Fast-Growing was the highest for the 2013 writing scaled score (M = 3850.31, SD = 559.62) and Major Urban had the lowest score (M = 3641.54, SD = 537.82).

As a result of significant F-ratios on the Welch F-test for 2013 reading scaled score, 2013 math scaled score, and 2013 writing scaled score, the Scheffe post hoc test was performed. This test was selected because it was conservative, and it corrected for alpha when used for simple and complex comparisons of students' performance on the reading, writing, and math 2013 STAAR tests.

Since the Welch's F-test was statistically significant for reading, writing, and math, the Scheffe post hoc test was performed to determine what community had the highest and lowest student scores on the 2013 reading, 2013 writing, and 2013 math STAAR tests. The Scheffe test revealed traditionally-certified and alternatively-certified

teachers' students achievement varied on the reading, writing, and math 2013 STAAR tests based on whether they were educated in an urban, rural, or suburban area.

As shown in Figure 1, the 2013 Charter (M = 1644.28, SD = 378.40) student reading STAAR test scores of alternatively-certified and traditionally-certified teachers were greater than the student scores of the alternatively and traditionally-certified teachers student scores in the following communities: Major Urban, (M = 1630.67, SD = 309.83) and Other Central City (M = 1641.74, SD = 286.37).

Major Suburban, (M = 1655.43, SD =289.36), Non-Metro Fast-Growing (M =1682.29, SD= 275.62), Non-Metro Stable (M =1644.85, SD = 299.97), and Other C. C. Suburban (M = 1652.26, SD = 283.73) had greater group means in comparison to Charter. This indicated the student scores on the 2013 reading STAAR test scores of traditionally-certified and alternatively-certified teachers were higher when comparing to Charter 2013 reading STAAR test student scores of alternatively and traditionally-certified teachers.

Independent Town (M = 1640, SD = 299.70) had a lower group mean in comparison to Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Other C.C. Suburban (M = 1652.26, SD = 283.73), Rural (M = 1667.77, SD= 331.58), and Charter (M = 1644.28, SD = 278.40). Major Suburban (M = 1655.43, SD = 289.36) had a lower group mean in comparison to Non-Metro Fast-Growing (M =1682.29, SD = 275.62), Rural (M = 1667.77, SD = 331.58), and Independent Town (M = 1640.69, SD = 299.70). Major Suburban had a greater group mean in comparison to the following: Non-Metro Stable (M = 1644.85, SD = 299.97), Other Central City (M = 1641.74, SD = 286.37), Charters (M = 1644.28, SD = 278.40), and Independent Town (M = 1640, SD = 299.70).

Major Urban (M = 1630.67, SD = 309.831) group mean was same in comparison to the following: Major Suburban (M = 1655.43, SD = 289.36), Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Non-Metro Stable (M = 1644.85, SD = 299.97), Other C.C. Suburban (M = 1652.26, SD = 283.73), Other Central City M = 1641.74, SD = 286.37), Rural (M =1667.77, SD = 331.58), Charter (M = 1644.28, SD = 278.40), Independent Town (M = 1640.69, SD = 299.70) and Major Suburban, (M = 1655.43, SD = 289.36).

The group mean of Non-Metro Fast-Growing (M = 1682.29, SD = 275.62) was greater than that of the following communities: Major Urban (M = 1630.67, SD = 309.83), Non-Metro Stable (M =1644.85, SD = 299.97), Other C.C. Suburban (M = 1652.26, SD = 283.73), Other Central Cities (M = 1641.74, SD = 286.37), Major Suburban (M =1655.43, SD = 289.36).

Also, Major Urban (M = 1630.67, SD = 309.83) had a lower group mean in comparison to Non-Metro Fast-Growing which indicated alternatively and traditionally-certified students scored lower on the reading 2013 STAAR test when compared to Non-Metro Fast-Growing alternatively and traditionally-certified teachers' student scores.

Non-Metro Stable (M = 1644.85, SD = 299.97) had a greater group mean in comparison to the following school community group means: Independent Town (M = 1617.73, SD = 312.72) and Major Urban (M = 1630.67, SD = 309.83). Rural (M = 1667.77, SD = 331.58), Non-Metro Fast-Growing (M = 1682.29, SD=275.62) had a greater group mean in comparison to Non-Metro Stable.

Other C.C. Suburban (M = 1652.26, SD = 283.73) had a greater group mean in comparison to Other Central City (M = 1641.74, SD = 286.37) and Major Urban (M = 1630.67, SD = 309.83. The following had a greater group mean than Other C.C.

Suburban: Non-Metro Fast-Growing (M = 1682.29, SD = 275.62), Rural (M = 1667.77, SD = 331.58), and Major Suburban (M = 1655.43, SD = 289.36).

As shown in Figure 2: the 2013 student math STAAR test scores of alternatively-certified teachers and alternatively-certified teachers of Independent Town (M =1617.73, SD = 312.72) had a greater group mean in comparison to the following communities group means: Other Central City (M = 1596.24, SD = 318.58), Charters (M = 1546.20, SD = 346.04), and Major Urban (M = 1576.80, SD = 343.77). Also, the following communities that had a greater group mean in comparison to Independent Town's group mean: Rural (M = 1646.97, SD = 350.22) and Non-Metro Fast-Growing (M = 1624.88, SD = 340.21).

Major Suburban (M = 1592.09, SD = 343.94) had a lower group mean in comparison to the following communities' group means: Non-Metro Fast-Growing (M = 1624.88, SD = 340.21) and Non-Metro Stable (M = 1618.48, SD = 321.26).

Major Urban (M = 1576.80, SD = 343.78) had a lower group mean in comparison to the following communities' group means: Other C. C. Suburban (M =1613.91, SD = 327.49), Other Central City (M = 1596.24, SD = 318.58), and Rural (M = 1646.97, SD = 350.22).

Non-Metro Fast-Growing (M = 1624.88, SD = 340.21) had a greater group mean in comparison to the following communities' group means: Charter (M = 1546.20, SD = 346.04), Major Urban (M = 1576.80, SD = 343.78), Other Central City (M = 1596.24, SD = 318.58), and Independent Town (M = 1617.73, SD = 312.72). The group mean for Rural (M = 1646.97, SD = 350.18) was greater than that of Non-Metro Fast-Growing.

Other C. C. Suburban (M = 1613.91, SD = 327.49) had a greater group mean that of the following communities: Charters (M = 1546.20, SD = 346.04) and Other Central City (M = 1596.24, SD = 318.58). Other C.C. Suburban 2013 math student STAAR test scores of alternatively-certified and traditionally-certified teachers were greater than the above student math scores of traditionally and alternatively-certified teachers. Rural student math STAAR test scores (M = 1646.97, SD = 350.18) had a greater group mean than Other C. C. Suburban 2013 student math STAAR test scores.

Other Central City (M = 1596.24, SD = 318.58) had a greater group mean in comparison to the group means of the following: Rural (M = 1646.97, SD = 350.18), Non-Metro Stable (M = 1618.48, SD = 321.26), and Independent Town (M = 1617.73, SD = 312.72). The following communities had a lower group mean than Other Central City: Charter (M = 1546.20, SD = 346.04), Major Urban (M = 1576.80, SD = 343.78), and Major Suburban (M = 1592.09, SD = 343.94).

The group mean for Rural (M = 1646.97, SD = 350.22) was greater than that of the following communities: Major Urban (M = 1576.80, SD = 343.78), Non-Metro Fast-Growing (M = 1624.88, SD = 340.21), Non-Metro Stable (M = 1618.48, SD = 321.26), Other C.C. Suburban, (M = 1613.91, SD = 327.49), and Other Central City (M = 1596.24, SD = 318.59).

As shown in Figure 3, the 2013 student writing STAAR test scores of alternatively-certified and traditionally-certified teachers of Charter (M = 3712.41, SD = 516.83) group mean was greater than the following communities' group means: Major Urban (M = 3641.54, SD = 537.82), Independent Town (M = 3643.84, SD = 520.51) and Non-Metro Stable (M = 3665.33, SD = 505.06). Non-Metro Fast-Growing (M = 3850.31,

SD = 559.62) and Major Suburban (*M*= 3778.89, *SD* = 558.31) had a greater group mean in comparison to Charter.

The 2013 student writing STAAR test scores of alternatively-certified and traditionally-certified teachers of Independent Towns (*M* = 3643.84, *SD* = 520.51) had a lower group mean in comparison to the group means of the following communities: Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M* = 3696.92, *SD* = 540.95), Rural (*M* = 3702.79, *SD* = 504.40), Major Urban (*M* = 3641.54, *SD* = 537.82) and Non-Metro Fast-Growing (*M* = 3850.31, *SD* = 559.62).

The 2013 student writing STAAR test scores of alternatively-certified and traditionally-certified teachers of Major Suburban (*M* = 3778.89, *SD* = 558.32) had a greater group mean than those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06), Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M* = 3696.92, *SD* = 540.95), Rural (*M* = 3702.79, *SD* = 504.40), and Charters (*M* = 3712.41, *SD* = 516.83).

Major Urban (*M* = 3641.54, *SD* = 537.82) had a lower group mean than those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06) and Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02). Rural (*M* = 3702.79, *SD* = 504.40), Charters (*M* = 3712.41, *SD* = 516.83), Independent Town (*M* = 3643.84, *SD* = 520.57), and Major Suburban (*M* = 3778.89, *SD* = 558.31).

Non-Metro Fast-Growing (*M* = 3850.31, *SD* = 559.62) had a greater group mean in comparison to those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06), Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M*

= 3696.92, *SD* = 540.95), Rural (*M* = 3702.79, *SD* = 504.40), and Charters (*M* = 3712.41, *SD* = 516.82).

Non-Metro Stable (*M* = 3665.22, *SD* = 505.06) had a lower group mean in comparison to those of the following communities: Other C. C. Suburban (*M* = 3729.64, *SD* = 537.02), Other Central City (*M* = 3696.92, *SD* = 540.95) and Rural (*M* = 3702.79, *SD* = 504.40).

Independent Town (*M* = 3643.84, *SD* = 520.57) and Major Urban (*M* = 3641.54, *SD* = 537.82) had group means lower than that of Non-Metro Stable (*M* = 3665.22, *SD* = 505.06).

Other C.C. Suburban (*M* = 3729.64, *SD* = 537.02) 2013 student writing STAAR test scores of alternatively and traditionally-certified teachers had a group mean greater than those of the following communities: Non-Metro Stable (*M* = 3665.22, *SD* = 505.06), Other Central City (*M* = 3696.92, *SD* = 540.95), and Independent Town (*M* = 3778.89, *SD* = 558.31).

Other Central City (*M* = 3696.92, *SD* = 540.95) had a greater group mean in comparison to Major Urban (*M* = 3641.54, *SD* = 537.82).

Discussion of Results

The ANCOVA statistical test was used to determine if there were statistically significant differences in student performances on the STAAR tests when taught by traditionally-certified and alternatively-certified teachers. As shown in Table 4, there were no significant differences in the effect of certification pathway on students' performance on grade 4 writing test when controlling for prior achievement. Also shown in Table 4 are grade 7 writing scores, which indicated that there were no significant

differences in the effect of certification pathway on students' performance when controlling for prior achievement. The student scores on the 4th grade STAAR writing tests of traditionally-certified and alternatively-certified teachers were not different when controlling for prior achievement. The student scores on the 7th grade STAAR writing tests of traditionally-certified and alternatively- certified teachers were not different when controlling for prior achievement. A reason why there were no differences in the student scores for the two teacher types could be that the teachers selected to teach these two tested grades for writing may have been veteran teachers. It is not known if principals placed their best teachers in the tested grades.

The Levene's Test for Equality of Variances, as shown in Table 3, for Grades 4 and 7 were found not to be violated for the analyses with a $p > .05$. The Levene's Test is non-significant, which means that the variances are equal. The ANCOVA statistical tests for 3rd, 5th, and 7th reading scores revealed that there were no significant differences in the effect of certification route on students' performance on the reading tests when controlling for prior achievement. However, the ANCOVA results for Grades 4, 6, and 8 reading scores revealed that there were significant differences regarding the effect of certification route on students' performance on the reading test of traditionally-certified and alternatively-certified teachers when controlling for prior achievement. The mean scale scores of traditionally-certified teachers were significantly, but only slightly higher than alternatively-certified teachers' mean scaled scores on the reading STAAR tests. The effect sizes for 3rd and 7th grade reading scores were $\acute{\eta}p^2=.001$, very small as shown in Table 7. The 3rd, 6th, and 7th grade Levene's Test for Equality of Variances results indicated that variances were different for the reading scaled scores. Ultimately, the

homogeneity of variances was violated, as shown in Table 8. When sample sizes are large, small differences in group variances can produce a significant Levene's test.

The potential reasons for no statistically significant differences in the 3rd, 5th and 7th grade student reading scores of traditionally and alternatively-certified teachers are the following: First, the students are tested six consecutive years on reading skills and teachers are able to fill any academic deficiencies from one grade level to the next depending on student needs. Second, school systems are aware of the importance of student test scores and teachers may have received professional development on reading instructional strategies since reading is tested in 3rd, 4th, 5th, 6th, 7th, and 8th grades.

The ANCOVA math scores for 3rd and 8th grade math revealed that there were no significant differences in the effect of certification routes on students' performance on the math STAAR tests when controlling for prior achievement. The effect size for 3rd grade math was .18. The student scores on the 3rd and 8th grade STAAR math tests of traditionally-certified and alternatively-certified teachers were not different based on the teachers' certification route when controlling for prior achievement. A reason for the math scores of the traditionally-certified and alternatively-certified teachers not being statistically significant is the teachers may had the opportunity to receive training or professional development on the math skills tested on the STAAR math test or possess more expertise in the content area of mathematics. The ANCOVA analyses for 4th grade, 5th grade, 6th grade, and 7th grade math scores revealed that there were significantly, but only slightly differences in the effect of certification routes on students' performance on the math STAAR test when controlling for prior achievement, as shown in Table 11. The mean scale scores of traditionally-certified teachers were significantly, but only

slightly higher than the mean scale scores of alternatively-certified teachers on the 4th and 5th grade STAAR math tests. The mean scale scores of alternatively-certified teachers were significantly, but only slightly higher than the mean scale scores of traditionally-certified teachers on the 6th and 7th grade STAAR math tests. The effect size for 6th grade math was $\acute{\eta}p^2$=.001, very small as shown in Table 11.

Grades 5, 6, and 7 math scores violated the assumptions of homogeneity of variances, as shown in Table 12. When sample sizes are large, small differences in group variances can produce a significant Levene's test.

The one-way Analysis of Variance test (ANOVA) was performed to determine whether or not there is a measurable difference between the writing, reading, and math scores of students in classrooms taught by alternatively-certified teachers and the scores of students in classrooms taught by traditionally-certified teachers in rural, urban, and suburban schools in Grades 3 through 8. The Levene's Test for Equality of Variances as shown in Table 14 was violated, demonstrating that significant differences between the variances exist. The one-way ANOVA was performed, as shown in Table 15. However, due to the violation of the homogeneity of variance, a more robust test was performed to adjust for groups with larger sample sizes and large variances. The adjusted *F*-ratio from the Welch *F*-test must be used to determine whether or not group differences were statistically significant and if a post hoc test was needed, as shown in Table 16. The Welch *F*-ratio revealed statistically significant results. The Welch *F*-ratio results were statistically significant at $p < .05$ for reading scaled scores, math scaled scores, and writing scaled scores; therefore, the Scheffe post hoc test was performed. This test was selected because it was conservative and because it corrected for alpha when used for

complex comparisons, as shown in Table 17. The Scheffe post hoc test revealed that students' scores of traditionally-certified teachers were significantly higher than the scores of students with alternatively-certified teachers.

When analyzing the Scheffe post hoc test, several interesting findings about the different areas were revealed. These included the following: the group mean for the Major Urban group was lower when compared to Major Suburban, Non-Metro Fast-Growing, Non-Metro Stable, Other C.C. Suburban, Other Central City, Rural, Charter, and Independent Town in reading, math, and writing, as shown in Table 17. Major Urban 2013 reading, math, and writing scores of traditionally-certified and alternatively-certified teachers were lower than all the other areas. This could be attributed to teachers having fewer years of teaching experience, the amount of professional development the teachers received, or whether or not they have certified teachers in the tested content areas. Urban school divisions are more likely to approve provisional teaching licenses to prospective teachers which could have an effect on student performance on state standardized tests because teachers are not experts in their content areas. Students in Urban school divisions may have less support at home due to parents working multiple jobs which may result in the lack of parental assistance with helping students with homework and school assignments. Poverty and the lack of sufficient funding from the locality are issues that many Urban school districts are confronted with that have a direct effect on the allocation of instructional resources available to the school divisions and the teachers. Also, rural areas had a greater group mean in comparison to Major Urban, Non-Metro Fast-Growing, Non-Metro Stable, Other C. C. Suburban, and Other Central City in math. The Rural student math scores of traditionally-certified teachers and alternatively-

certified teachers were higher probably due to veteran teachers with experience teaching math and also due to a more stable faculty. The Major Suburban group had a lower group mean in the content area of reading in comparison to Non-Metro Fast-Growing, Rural, and Independent Town. Major Suburban reading students' scores of traditionally-certified and alternatively-certified teachers were lower than the socioeconomic demographics alternative and traditionally-certified teachers; scores listed above. The difference in the scores may be contributed to a decrease or absence of professional development opportunities in the area of reading. In the content area of math, Major Suburban had a lower group mean in comparison to the following communities' group means: Non-Metro Fast-Growing, and Non-Metro Stable. Major Suburban student math scores of traditionally-certified and alternatively-certified teachers were lower than socioeconomic demographics alternatively and traditionally-certified teachers' student scores for the above listed communities. As mentioned before, there could have been less professional development available for math teachers in the Major Suburban areas. An area which had higher scores in the content area of writing, Major Suburban, had greater group mean in comparison to Non-Metro Stable, Other C.C. Suburban, Other Central City, Rural, and Charters, as shown in Table 17. This difference in the student scores could be attributed to a lower transience rate for teachers and students. Major Suburban writing student scores of traditionally and alternatively-certified teachers were higher than the other community area's alternatively-certified and traditionally-certified student scores.

Possible Explanations for Findings

The results from the ANCOVA analyses revealed that there were significant differences regarding the effect of certification route on the reading scores of students taught by traditionally-certified and alternatively-certified teachers on the 4th, 6th, and 8th grade reading tests. Traditionally-certified teachers' student scores were significantly, but only slightly higher than the alternatively-certified student scores on the 4th, 6th, and 8th grade STAAR reading tests. The effect sizes for these three reading tests were so small, essentially non-existent. A possible explanation for the findings is that the students were not familiar with all of the reading skills from the 3rd grade STAAR test and alternatively-certified teachers were not equally equipped with the expertise needed to get their students to master the reading skills tested from 4th, 6th, and 8th grades. Social promotion could also be an external factor for the differences in the 4th grade student reading scores of alternatively-certified teachers.

The results from the ANCOVA analyses also revealed that there were significantly, but only slightly differences regarding the effect of certification route on the math scores of students taught by traditionally-certified and alternatively-certified teachers on grades 4 through 7 math tests. The effect sizes were so small, essentially non-existent. grade 3 math STAAR test revealed that there were no significant differences in the 3rd grade student scores of traditionally-certified and alternatively-certified teachers when controlling for prior achievement; however, the effect size was greater than .13. This could be attributed to the fact that both of the teacher types received professional development on 3rd grade math skills or more veteran teachers with more classroom experience in teaching 3rd grade math impacted the student scores on the

3rd grade STAAR math test. The 8th grade math test revealed that there were no statistically significant differences in the effect of certification route on the math scores of students taught by traditionally-certified and alternatively- certified teachers while controlling for prior achievement. A possible explanation is the students are accustomed to taking the STAAR test so they are familiar with the testing process and the teachers who teach 8th grade have a higher level of expertise since the 8th STAAR test is middle school math. Many veteran teachers in school systems often teach the higher-level math classes because their certification usually allows them to teach high school courses as well. Traditionally-certified teachers had significantly, but only slightly higher scores than alternatively-certified teachers' student scores on the 4th and 5th grade STAAR math tests. This could be attributed to the fact that the teachers may have more experience and the students have mastered the mathematical skills necessary to achieve in 4th and 5th grade mathematics. Alternatively-certified teachers' student scores were significantly, but only slightly higher than traditionally-certified teachers' student scores on the 6th and 7th grade STAAR math tests. The alternatively-certified teachers may have had the opportunity to teach 6th and 7th grade mathematics for several years and are competent in teaching the higher level mathematics. The students of the alternatively-certified teachers may have mastered the previous grade level math skills that enabled them to build on their prior knowledge.

In 4th and 7th grade writing, there were no statistically significant differences in the student scores of traditionally-certified and alternatively-certified teachers while controlling for prior achievement. The teachers in 4th and 7th grade writing may have received the same professional development and may be veteran teachers.

In the Blackburn et al. (2006) study, the characteristics of a typical alternative teacher preparation program were found to include an internship and mentorship by skilled professionals. Most alternative teacher candidates have already earned a bachelor's degree and have knowledge in the content area based on work experience, but lack the educational coursework. The statistical differences shown in the ANCOVA analyses could be attributed to the following: the lack of experience of the teachers, inability to effectively engage the students, the lack of knowledge about educational pedagogy necessary to reach all diverse learners in the classroom, and the absence of having veteran teachers in the content areas, whether they are alternatively-certified or traditionally-certified. According to Goldhaber, Liddle, and Theobald (2013), after three to five years of classroom experience, there is no difference in student performance of alternatively-certified and traditionally-certified teachers.

The ANOVA analyses revealed what was expected about the Major Urban group mean scores being lower than Major Suburban and Rural socioeconomic areas. In Urban schools, teachers often have less teaching experience and students tend to be more transient. Once teachers receive tenure and/or three years of experience, they may tend to move to a school division that pays more money, the students are less transient, and the students have a higher socioeconomic status. Students from urban areas are often more prone to have outside poverty factors that affect their learning in the classroom, and teachers have to work harder to help their students overcome obstacles and teach them to succeed while in class. This is especially important for teachers because student performance is now a component of the teachers' overall evaluation in some states.

Implications

These findings have implications for the local, state, and national levels. First, it is necessary for states to ensure the quality and effectiveness of alternative certification programs because many teachers pursue this route in order to become a certified teacher. In this study, 66.71% of the teachers were certified by an alternative certification program. The research has shown that with professional development, mentorship, and an internship, alternatively-certified teachers can be equally effective at engaging students and increasing student performance on state standardized tests (Bol, Gimbert, & Wallace, 2007). This study suggests that some statistical differences regarding the effect of certification routes exists when comparing students' scores of traditionally-certified and alternatively-certified teachers when controlling for prior achievement in some grades for reading and math. This study showed small effect sizes were less than .02, which is miniscule in light of the ANCOVA analyses. This confirms the literature indicating that alternatively-certified teachers' student performance on standardized tests is comparable to traditionally-students' performance on standardized tests after three to five years of teaching experience (Goldhaber, Liddle, & Theobald, 2013).

Second, this study can benefit local, state, and national government agencies that are responsible for monitoring, implementing change, and evaluating traditional certification programs and alternative certification programs. The large sample size, 2nd largest to date, helped to validate the previous research because it resulted in analysis of a large sample of traditionally and alternatively-certified teachers' student scores on the STAAR tests using ANCOVA analyses. Last, this study can provide information on whether student promotion to the next grade level should be based solely on high-stakes

testing such as the STAAR test and other standardized testing that is used to evaluate students' mastery of concepts taught in the classroom over a long period of time with the expectation that every child pass the test.

Future Research

Given the significant findings in this study, there are five recommendations for further research on this subject. First, include a qualitative component that involves capturing insight from teachers that focuses on their experiences in regards to their certification programs. This should ideally include positive and negative perspectives on this licensure track. Collecting feedback from teachers could benefit certification programs and school divisions that select prospective teachers for entrance into the teacher preparation programs and employment into the nation's school divisions. It would also allow educators to gain insight into how teachers are placed in testing grade levels. Their perspectives could be used to improve the quality and effectiveness of teacher preparation programs.

The second recommendation is to obtain qualitative feedback from students of traditionally-certified and alternatively-certified teachers. Their perspectives on whether teachers were able to engage them in the classroom and motivate them to achieve on standardized tests are important aspects of the teachers' ability to help their students succeed in the classroom.

The third recommendation is to survey building administrators to obtain their feedback on whether alternatively-certified teachers immersed easily into the culture of the school and were competent in using instructional strategies to engage students. Another category of questions that should be included within the survey are: were

teachers able to implement the professional development activities in their teaching and did they have a passion to become an expert in their content area.

The fourth recommendation is to survey universities' schools of education and alternative licensure programs to determine how successful preparation programs are producing highly-qualified teachers who improve student performance on state standardized tests. These teachers can serve as potential mentors to new teachers in the field of education.

The fifth recommendation is to conduct a similar study and research the various types of alternative teacher programs, certification routes teachers take, the type of certification obtained, the duration of certification process, and associate these variables with student achievement results.

The sixth recommendation is the collect feedback (effective vs. ineffective alternatively certified teachers-what helped or hindered their effectiveness) that could benefit school divisions and certification programs.

The final recommendation is to update data and identify trends when revisiting the study. Data could be collected to determine if years of experience have an effect on traditionally-certified and alternatively-certified student scores on state standardized tests. Colleges and universities could follow their graduates' students test scores to determine effectiveness compared to other college or university graduates' student test scores.

Conclusion

The results of this research supported the idea that significant differences exist between the student scores of traditionally-certified teachers and alternatively-certified teachers' student scores in reading and math in some grades. Findings from the Scheffe

post hoc test revealed that the group mean of the Major Urban communities was smaller in reading, math, and writing in comparison to Rural and Major Suburban communities. This means the math, reading, and writing scores of traditionally-certified and alternatively-certified teachers from Major Urban communities are lower than the student scores of traditionally-certified and alternatively-certified teachers' student scores in Rural and Major Suburban communities. This is not surprising because there is a lack of veteran teachers in many urban school divisions. In addition, sadly poor, urban students are often more transient than other school and student populations.

The findings of this study can be a precursor to the efforts of other states evaluating whether or not traditionally-certified teachers and alternatively-certified teachers students' scores on state standardized test are comparable when including other quantitative variables, such as years of teaching experience, teacher performance on basic skill tests, and teacher quality data, to ensure prospective teachers have the ability to provide the teaching quality needed to increase the probability of successful student achievement.

References

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *Future of Children ,17,*(1), 46.

Boyd, D., Grossman P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31,*(4), 416-440.

Blackstone, D. L. (2010). A quantitative correlation study of teacher preparation program on student achievement. Available from Proquest Dissertation and Theses database. Retrieved from www.odu./ database. University of Phoenix.

Brill, J. M. (2001). Situated Cognition. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. Retrieved from http://projects.coe.uga.edu/epltt/

Brinkman, B. (2014). *The impact of teacher preparation on academic achievement of eighth grade students in a south Texas school district*. (Doctoral Dissertation , Texas A.M. University-Corpus Christi). Retrieved from www.odu./database.

Clark-Tuttle, C., Anderson, T., & Glazerman, S. (2009). ABCTE teachers in Florida and their effect on student performance. *Mathematica Policy Research Inc.* 6215-400.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2007b). Teacher credentials and student achievement in high schools: *A cross-subject analysis with student fixed effects.* (CALDER Working Paper 11). Washington, DC: The Urban Institute. Retrieved from http://www.caldercenter.org./PDF/1001104 Teacher Credentials HighSchool.pdf.

Clotfelter, C. T., Ladd, H.F., & Vigdor, J.L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student effects. *The Journal of Human Resources, 3,*(45), 665-681.

Collins, A. (1988). Cognitive apprenticeship and instructional technology. (Technical Report No. 6899). BBN Labs Inc., Cambridge, MA.

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., Deke, J., & Warner, E. (2009). *An evaluation of teachers trained through different routes to certification, Final Report.* (NCEE 2009-4043). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/pubs/20094043/index.asp

Cordova, R. (2001). *A formative evaluation of the Texas teacher licensure system.* (Doctoral Dissertation, Southwest Texas State University). Retrieved from www.odu/database.

Croninger, R. G., Rice, J. K., Rathbun, A., & Nishio, M. (2003). Teacher qualifications and first grade achievement: A multilevel analysis. *Center for Education Policy and Leadership.* Document OP-03-2. Retrieved from http://www.education.umd EDPS/CEPAL/pdf/oc-papers/CEPAL-OP-)03-2.pdf

Education Commission of the States. (2002). No state left behind: *The challenge and opportunities of ESEA* 2001. Denver, CO: Author, 30.

Dee, T. & Jacob (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30,*(3), 418-420.

DePoy, E. & Gitlin (2011). Experimental-type Design, 4th Ed., Introduction to research:

Understanding and applying multiple strategies. 114-115. St. Louis, MO: Elsevier Mosby.

Evans, L. (2010). Professionals or technicians? Teacher preparation programs and occupational understandings. *Teachers and Teaching: Theory in Practice, 16,* (2). 184-202.

Feistritzer, C. (2008). What are alternate routes to teacher certification? E. Feistritzer (Ed.), *Alternate Routes to Teaching* (2-8; 65-87). Upper Saddle River, New Jersey: Pearson Education, Inc.

Ferguson, R. (1998). Teachers' perception and expectations and the Black-White test score gap. In C. Jencks and M. Phillips (Eds.), *The Black-White Test Score Gap* (273-317). Washington, DC: Brookings Institute Press.

Field, A. (2011). Discovering Statistics Using SPSS. Analysis of Covariance. ANCOVA 395-400. Singapore: SAGE Publications Asia-Pacific Pte Ltd.

Fraser, J. (2007). Preparing America's Teachers: A History. Teachers for Cities, Teachers for Immigrants. 160-161. New York: Teachers College Press.

Gansle, K. A., Noell, G.H., & Burns, J.M. (2012). Do student achievement outcomes differ across teacher preparation programs? An analysis of teacher education in Louisiana. *Journal of Education, 63,*(5), 304-317.

Gimbert, B., Bol, L., & Wallace, D. (2007). The influence of teacher preparation on Student achievement and the application of national standards by teachers of mathematics in urban secondary schools. *Education and Urban Society, 40,*(49), 95-104.

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: *The*

impacts of Teach for America on student achievement and other outcomes. Journal of Policy Analysis and Management, 25(1), 80-83.

Goldhaber, D. & Brewer, D. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Education Evaluation and Policy Analysis, 22,* (4), 139-140.

Goldhaber, D., Liddle, S., Theobald, R. (2013). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Economics of Education Review, 34,*29-44.

Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education, 61,*(35), 37-40.

Hammond, L., Holtzman, D., Gatlin, S., & Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13,* (42), 3-20.

Heilig, J. & Jez, S. (2010). Teach for America: A Review of the Evidence. Tempe, AZ: *Education and the Public Interest Center & Education Policy Research Unit* Retrieved from http://epicpolicy.org/publication/teach-for-america.

Hershberg, T. & Robertson-Kraft, C. (2010). Maximizing the opportunity provided by Race to the Top. *Perspective on Urban Education*, Summer, 128-129.

Kaplan, L. & Owings, W. (2011). *American Education: Building a Common Foundation.* 44-47. Belmont, CA: Wadsworth Cengage Learning.

Kane, T., Rockoff, J., & Staiger. (2007). Photo finish: Certification doesn't guarantee a winner. *Education Next, 7,*(1), 60-67.

Kane, T., Rockoff, J., & Staiger (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27,* 616-620.

Kee, A. (2012). Feelings of preparedness among alternatively certified teachers: What is the role of program features? *Journal of Teacher Education, 63,*(23), 23.

Lackzo-Kerr, I. & Berliner, D. (2002). The effectiveness of "Teach for America" and Othe runder-certified teachers on student academic achievement: A case of harmful public policy." *Education Policy Analysis Archives, 10,*(3), 25-41.

Leedy, P. & Ormrod, J. E. (2005). Experimental and Ex Post Facto Designs, 8th Ed. Practical research: Planning and design. 223-225. Upper Saddle River, New Jersey: Pearson Education, Inc.

Lindquist E. F. (2001). The Iowa testing programs: A retrospective view. *Education, 91,*(1), 10.

Marszalek, J.M. Odom, A. L., LaNasa, S.M., & Adler, S. A. (2010). Distortion or clarification: Defining highly qualified teachers and the relationship between certification and achievement. *Educational Policy Archives, 18,* 3-6.

Meyers, S. (2011). Life skills training through situated learning experiences: An alternative instructional model. *International Journal of Special Education, 26,*(3), 142-149.

McGuinn, P. (2012). Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy, 26,*(1), 137-138.

Mitchell, K.J., Robinson, D.Z., Plake, B.S., & Knowles, K.T. (2001). Testing and

Licensing Beginning Teachers. *Testing teacher candidates: The role of licensure tests in Improving teacher quality, 45*. Washington, D.C.: National Academy Press.

Nelson, S., McGhee, M., Meno, L., & Slater, C. (2007). Fulfilling the promise of educational accountability. *Phi Delta Kappa, 88* (9), 702-709.

Nichols, S. L., Glass, G.V., & Berliner, D. C. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Educational Policy Analysis Archives, 20,*(20), 2-4.

Nunnery, J., Kaplan, L., Owings, W., & Pribesh, S. (2009). The effects of Troops to Teachers on student achievement: One state's study. *NASSP Bulletin, 93,*(4), 259-262.

Owings, W., Kaplan, L., Nunnery, J., Marzano, R., Myran, S., & Blackburn, D. (2006). Teacher quality and Troops to Teachers: A national study with implications for principals. *NASSPBulletin ,90,*(2), 103-105.

Papay, J., West, M., Fullerton, J., & Kane, T. (2012). Does an urban teacher residency Increase student achievement? Early evidence from Boston. *Educational Evaluation and Policy Analysis ,34,*(4), 413-414.

Peabody, Z. (2003, December). Paige defends Houston Independent School Division progress-Critics accused of undermining reform. *Houston Chronicle.*

Raymond, M., Fletcher, S.H., & Luque, J. (2001). Teach for America: An evaluation of teacher differences and student outcomes in Houston, Texas. Stanford, CA: The Hoover Institution, Center for Research on Educational Outcomes.

Rudalevige, A. (2003). The politics of No Child Left Behind. *Education Next, 64,*

Texas Education Agency. Retrieved from http://ww.tea.state.tx.us

Sass, T.R. (2011). Certification requirements and teacher quality: A comparison of Alternative routes to teaching, (CALDER *working paper* 64). Retrieved from files.eric.ed/gov/fulltext/ED529179.pdf

Sprinthall, R. C. (2012). Analysis of Variance. Hartman, S., 9th Ed. Basic Statistical Analysis, 349-350. Boston, MA: Allyn & Bacon.

Sprinthall, R. C. (2012). The Fundamentals of Research Methodology. Hartman, S., 9th Ed. Basic Statistical Analysis (pp. 224-225). Boston, MA: Allyn & Bacon.

Strauss, R. P. & Sawyer, E. A. (1986). Some new evidence on teacher and student competencies. *Education Review, 5*,(1), 41-48.

U.S. Department of Education. (2009). Race to the Top: Program executive summary (Technical Report). Washington DC: Author. Retrieved from https://www.2edgov/programs/racetothetop/executive-summary.pdf

Veltri, B. T. (2012). Teaching for America: It's more about leading than teaching. *Educational Leadership, 69*,(8), 62-69.

Walsh, K. & Jacobs, S. (2007). Alternative certification isn't alternative. *National Council on Teacher Quality.* Retrieved from http://news.heartland.org/ sites/all/modules/custom/heartland_migration/files/pdfs/22264.pdf

Wong, K. & Nicotera, A. (2007). The context of accountability. In James Guthrie (Ed.), *Successful schools and educational accountability.*(p.26). Boston, MA: Pearson.

Xu, Z., Hannaway, J., & Taylor, C. (2011). Making a difference? The effects of Teach for America in high school. *Journal of Policy Analysis and Management, 30*,(3), 447-469.

Zimmerman, B. J. & Dibenedetto, M. K. (2008). Mastery learning and assessment: Implications for students and teachers in an era of high-stakes testing. *Psychology in the Schools, 45,*(3), 206-215.

VITA

Patrizia Antoinette Grigsby
Department of Educational Foundations and Leadership
Darden College of Education
Old Dominion University Norfolk, VA. 23529
pgrig001@odu.edu

Academic Degrees

Ph.D. Educational Leadership, December, 2015, Old Dominion University
Ed. S. Administration and Supervision, December 2008, Old Dominion University
M.A. Curriculum and Instruction, May 2004, University of Saint Mary's
B.S. Medical Technology, June 1990, University of Tennessee

Professional Experience

Science Program Specialist
August 2008 – Present
K-6 Instructional Specialist Portsmouth Public Schools
Provide instructional support to teachers and administrators in K-6 schools

Teacher
August 1997 – May 2008
Various School Divisions