2018

Impact of Educator Preparation Routes and Teacher Characteristics on Student Academic Achievement

Robert Hagerman
hagerman@marshall.edu

IMPACT OF EDUCATOR PREPARATION ROUTES AND TEACHER CHARACTERISTICS ON STUDENT ACADEMIC ACHIEVEMENT

A dissertation submitted to
the Graduate College of
Marshall University
In partial fulfillment of
the requirements for the degree of
Doctor of Education
in
Curriculum and Instruction
by
Robert Hagerman
Approved by
Dr. Samuel Securro, Committee Chairperson
Dr. Lisa Heaton
Dr. Andy Whisman

Marshall University
August 2018

APPROVAL OF DISSERTATION

We, the faculty supervising the work of **Robert Hagerman**, affirm that the dissertation, *Impact of Educator Preparation Routes and Teacher Characteristics on Student Academic Achievement*, meets the high academic standards for original scholarship and creative work established by the EdD Program in **Curriculum and Instruction** and the College of Education and Professional Development. This work also conforms to the editorial standards of our discipline and the Graduate College of Marshall University. With our signatures, we approve the manuscript for publication.

Sam Securro Digitally signed by Sam Securro Date: 2018.08.06 10:31:36 -04'00'	Dr. Sam Securro	08/6/2018
Dr. Sam Securro Curriculum and Instruction	Committee Chairperson Major	Date
Lisa Heaton Digitally signed by Lisa Heaton Date: 2018.08.06 10:31:18 -04'00'	Dr. Lisa Heaton	08/6/2018
Dr. Lisa Heaton Curriculum and Instruction	Committee Member Major	Date
	Dr. Andrew Whisman	08/6/2018
Dr. Andrew Whisman	Committee Member	Date

DEDICATION

I dedicate this dissertation to my wife Robyn and to my children Hailey and Kyle. I could not have done this without them. To my wife, for not letting me give up and for her unconditional support. To my children, for inspiring me and believing in me. Thank you for making this possible.

ACKNOWLEDGMENT

I would like to give very special thanks to Dr. Securro for sticking with me and pushing me, and for his dedication to my success. You are an incredibly patient individual. To my Committee members Dr. Heaton and Dr. Whisman, thank you for your support, knowledge, willingness to assist me through this process, and for your kindness. In addition, I would like to thank the many faculty members at Marshall University for their guidance and my work colleagues who never hesitated to guide me no matter what and who have been a constant source of encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This study investigated the effect of traditional and alternative teacher preparation, years of service, and type of licensure held and teacher quality on English language arts and mathematics summative assessment scaled scores and performance levels among West Virginia students in grades 5 and 11. Specifically, this research analyzed the existing assessment data in West Virginia's 55 counties, regarding teacher preparation routes, teacher experience, teacher licensure, and the teacher's Highly Qualified (HQ) designation to determine the effect on student achievement. The study was designed with the aforementioned variables and applied a standardized summative content assessment outcome to the two grade levels and two content disciplines. Data analysis indicated that the majority of teacher quality variables had a statistically significant impact on student achievement. As with every aspect of education, various socio-economic variables and teacher and student characteristics not measured in this study, and not known, may impact the standardized achievement results of the students.

CHAPTER 1

INTRODUCTION

Since the beginning of formal public education in the 1600s, there has been an ongoing debate arguing the best way to prepare teachers to positively impact student outcomes (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2008). Education constituencies, including teachers, administrators, higher education faculty, and policy makers, have advocated for teacher preparation that includes more time in the classroom as candidates learn the content they will teach. Others within these constituencies believe individuals who have mastered the content knowledge and have succeeded in professions other than education are better poised to serve P-12 students as more effective teachers. Many of these same individuals profess that a field-based program such as alternative education where candidates are in the classroom as the teacher of record from day one is more effective. As a result of teacher shortages and critical needs, states offer a variety of pathways and often define these very differently (Mader, 2013). Some of these pathways continue to exist while others, such as the *Transition to Teaching* program in West Virginia, are defunct.

According to a report by the Office of Innovation and Improvement at the U.S. Department of Education (2004), states have created alternative programs where that state claims the field should allow other professionals to enter the profession laterally and decrease the entrance requirements. The majority of states across the country have an alternate route for teacher certification compared to the traditional four-year route. As a result of the turn-over rate and lack of qualified applicants, local education agencies and school district administrators advocate to recruit professionals from industry and other fields, including military veterans, to become teachers. Because there is a large number of vacancies in schools across the country,

states are considering less restrictive and potentially less rigorous pathways to becoming a qualified teacher. Their opposing counterparts, institutional faculty and many practitioners, argue that more restrictions, requirements and rigor should be introduced to improve the quality that teacher education students receive (Walsh, Joseph, & Lewis, 2016).

The majority of the research on teacher quality within the context of educational policy is inconclusive (Rice, 2003) and has been conducted to investigate two aspects of the teacher's quality: either preparation or experience. These approaches limit the scope of the findings and place the main focus on elements of either the preparation process or general years of experience rather than on the outcome of such preparation as demonstrated by student achievement. Most studies have taken place in a single district, school or city, rather than in whole states or across multiple districts (Boyd et al., 2008; Robinson, 2011; Suell and Piotrowski, 2006;). The focus has been on the difference between alternative programs and traditional routes (Gimbert, Bol, & Wallace, 2007; Suell and Piotrowski, 2006). Through the years, available research examined the relationship between teachers prepared in alternative programs to teachers prepared in traditional programs. A main intent of the current study is to determine if the different preparation routes, and how teachers implement related teaching practices and standards, impact the academic achievement of students.

According to the No Child Left Behind Act (U.S Department of Education, 2004) all students must be taught by highly qualified educators. Within this mandate there are very specific measures of growth at the student and school levels. These growth measurements and benchmarks include student performance on the state-adopted assessments at a prescribed percentile or a pre-determined amount of student and school growth. These levels of performance are compared to previous years to determine subsequent growth. The resulting

measures of growth are directly related to funding and to the continuation of programs and initiatives. These initiatives include student support programs such as after school tutoring, funding for additional teachers for interventions, and instructional resources and technology to support school and student improvement.

Considering the federal landscape in recent years, states are under rigorous public and governmental scrutiny regarding the performance of their P-12 students and the effectiveness of their teachers measured by how students score on standardized achievement tests. In West Virginia, the state accountability system includes rigorous measures in order to produce performance data at both student and school levels. At the student level the measure is an annual standardized content test (math, science, English language arts, social studies) referred to as the West Virginia summative assessment. At the school level, the measure is the overall performance of the students in that school on the West Virginia summative assessment and other academic benchmarks for non-tested subject areas and the growth of these students through the years. Not meeting these benchmarks may result in sanctions that range from additional oversight by the state and the federal government to the loss of funding.

Schools in need of improvement (priority schools) must follow the US Department of Education Turnaround Principles. If these principles are not met, districts come under strict scrutiny and monitoring by the US Department of Education. With the new A-F grade accountability system, schools that receive an F grade for two consecutive years can potentially be taken over by the state. Without compliance, the state would potentially lose millions of dollars in federal monies that provide essential services to its students. These services include Title I, II, and III services and many other programs and initiatives. More states have shifted to new educator evaluation systems and many are tying these systems to teacher tenure, hiring

practices, and salaries (Springer et al., 2010). Colorado, Florida, Idaho, and Indiana have systems where teacher compensation is tied to performance (Behrstock-Sherratt & Potemski, 2013).

Because funding is instrumental for providing and supplementing many educational programs, states have adopted more stringent requirements when licensing teachers for their public schools. Federal funding is now requiring states to have rigorous systems in place addressing teacher effectiveness. These systems must ensure that teacher preparation programs adhere to strict accountability measures. Both school districts and institutions of higher education are required to produce data connecting teacher effectiveness to students' academic achievement. Many states are connecting new teachers and their performance to the institution from which they completed their preparation. For example, North Carolina, Louisiana, and Tennessee are among states that now have teacher preparation student performance data models. These models use student performance data to evaluate preparation programs. North Carolina matches student data to specific educator preparation programs and not just to the institution as a whole (NCTQ, April 2013)

In recent years, considerable research has been conducted in the area of teacher quality and student achievement. The majority of it has focused on specific programs or routes such as Teach for America and the Florida Alternative Program. Identifying high-quality and effective teachers continues to be a major goal for researchers as well as for policy makers throughout the country. When considering the same students, highly qualified and effective teachers have shown to increase student achievement from the 50th percentile to, in some cases, the 95^{th} percentile. Highly qualified teachers are defined as those who demonstrate content mastery via the state's allowable pathways. Effective teachers are defined as those whose students after a year of their instruction show growth. In addition, students who were taught for three years by high

performing teachers scored on average at the 96th percentile on summative assessments. Those who were taught by low-performing teachers for three years performed at the 44th percentile (Tucker & Stronge, 2005).

For the last 10 years, teacher accountability for student growth as required in the No Child Left Behind Act of 2001 has been a topic of great concern for states and the target of several initiatives such as Race to the Top designed to prompt states to implement education reforms to receive financial incentives. Some of the available funding for states and districts to provide student-related services and programs is now linked to teacher performance and student progress (No Child Left Behind Act and Teacher Accountability, FindLaw, 2009). Millions of dollars have been infused into national and local educational systems attempting to define, capture, and measure student growth. Student growth is commonly measured by how students score on summative standardized achievement tests implemented statewide from year to year. If a student achieves higher on the assessment on subsequent years it is an indication of growth. Often, how well students score on such standardized assessments is perceived to be a result of how effectively the teachers perform.

Based on the accountability measures West Virginia can receive funding for areas such as School Improvement Grants (SIG). To meet legislative mandates, it is imperative to examine the perceived impact on students' standardized test scores related to teacher preparation, years of service, type of licensure held, and teacher quality. Depending on the effect these variables might have on student achievement, West Virginia's ability to implement educational reform in these areas could significantly impact student achievement outcomes.

In order to enhance teacher preparation programs, shape policy regarding licensure, and determine the needed changes in teacher induction models currently in place, states, including

West Virginia, should investigate the relationship between these variables. Access to some funding streams is linked to how students perform. Some areas of the country tie teacher performance evaluations and merit pay to student performance and student growth. Therefore, it is critical to identify factors that will have a positive impact on these measures (Rosales, 2014; Springer et al., 2010). In order to inform the policy-making process and implement any needed changes and/or scale-up local initiatives, it is important to determine what will influence teacher performance and the quality of teaching as it relates to student outcomes.

There are a variety of methods and pathways available to certify public school educators. An understanding of preparation routes and methods and the impact teachers prepared under each have on the achievement of public school students in West Virginia can help identify best practices for student success.

Statement of the Problem

The purpose of the study is to determine the effect of teacher preparation, years of service, type of teacher licensure and teacher quality, on English language arts and mathematics West Virginia summative assessment scaled scores and performance levels (Levels l, 2, 3, and Level 4) in grades 5 and 11. Specifically, this research will analyze the existing assessment data in West Virginia's 55 counties, regarding teacher experience and the types of license held including initial Professional Licenses and Alternative licenses. What might be the effect, if any, on student academic achievement that may be moderated by these variables?

Research Questions

1. What are the differences between the West Virginia summative assessment scaled scores in English language arts and mathematics among West Virginia students in grades 5 and 11, taught by teachers trained in traditional and alternative teacher preparation programs?

2. What are the differences in categorical rankings (levels 1, 2, 3, and 4) on the West Virginia summative assessment in English language arts and mathematics among West Virginia students in grades 5 and 11 taught by teachers trained in traditional and alternative teacher preparation programs?
3. What is the effect of teacher experience on West Virginia summative scaled scores in English language arts and mathematics among West Virginia students in grades 5 and 11?
4. What is the effect of teacher experience on West Virginia summative assessment categorical rankings (levels 1, 2, 3, and 4) for mathematics and English language arts, among students in grades 5 and 11?
5. What is the effect of the type of teacher licensure (certified/noncertified) on West Virginia summative scaled scores in English and mathematics among West Virginia students in grades 5 And 11?
6. What is the effect of type of licensure (certified/noncertified) on categorical rankings in math and English, among students in grades 5 and 11?
7. What is the effect on math and English scaled scores by 5th grade math and English teachers with highly qualified status or non-highly qualified status?
8. What is the effect of highly and non-highly qualified teacher status on categorical rankings in English and mathematics among students in grades 5 and 11?

Null Hypotheses

1. There are no differences in West Virginia summative assessment scaled scores in English language arts and mathematics and related categorical rankings (levels 1, 2, 3 and 4) among West Virginia students in grades 5 and 11 taught by highly qualified teachers compared to their peers taught by non-highly qualified teachers.

2. There are no differences in West Virginia summative assessment scaled scores in English language arts and mathematics and related categorical rankings (levels 1, 2, 3 and 4) among West Virginia students in grades 5 and 11 taught by traditionally prepared teachers compared to their peers taught by alternatively prepared teachers.
3. There are no differences in West Virginia summative assessment scaled scores in English language arts and mathematics and related categorical rankings (levels 1, 2, 3 and 4) among West Virginia students in grades 5 and 11 taught by beginning teachers compared to their peers taught by experienced teachers.
4. There are no differences in West Virginia summative assessment scaled scores in English language arts and mathematics and related categorical rankings (levels 1, 2, 3 and 4) among West Virginia students in grades 5 and 11 taught by teachers holding a Professional Teaching Certificate compared to their peers who are noncertified teachers.

Operational Definitions

Highly Qualified Teacher.

a. A teacher who holds a bachelor's degree or higher and meets state certification requirements, including those certified through a West Virginia Board of Education (WVBE) approved alternative certification program. The teacher has an endorsement(s) in the core academic subject(s) and has successfully passed the state competency test in the content area.

b. A teacher is also Highly Qualified who has a minimum of a bachelor's degree with an academic major or advanced credential(s) in the core subject taught.

c. A teacher is also Highly Qualified who satisfied West Virginia's Highly Objective Uniform State Standard of Evaluation (HOUSSE), an optional method of documenting subject matter competency in a core academic subject(s) via classroom observations by the school

administrator and provided the individual previously held the appropriate license to deliver instruction in the core academic subject.

Non-Highly Qualified Teacher.

a. A teacher who holds a bachelor's degree or higher and who is not yet fully certified. A teacher is also non-highly qualified if he/she holds a Professional Teaching Certificate endorsed in the appropriate content area but who has not yet passed the state competency test –Praxis II in the content area.

b. A teacher who holds full certification endorsed in the appropriate content areas but does not have an academic major or advanced credential in the subject taught and who has not satisfied the West Virginia's Highly Objective Uniform State Standard of Evaluation (HOUSSE) definition.

Traditionally Prepared Teacher – A teacher who has completed a traditional baccalaureate preparation program at a West Virginia institution of higher education, and who has completed a student teaching experience and who has passed all West Virginia required Praxis exams and holds full certification in the appropriate content area.

Alternatively Prepared Teacher – A teacher with a baccalaureate degree who has completed an alternative program, and who completes student teaching on the job and who is employed as a teacher while completing the program.

Beginning Teacher – A teacher who has between 0 and 3 years of full-time teaching experience in the core subject taught.

Experienced Teacher – A teacher who has a combined total of 5 or more years of full-time teaching experience in the core subject taught.

Teaching License – The license issued by the West Virginia Board of Education held by the teacher. These are:

Initial License – The first license valid for a period of three years issued to a teacher who meets all licensure requirements in West Virginia.

Five-Year license – A license issued to a teacher who has a minimum of two years of teaching experience and who has completed a beginning teacher internship.

Temporary Teaching License – A non-renewable, one-year license issued to a teacher who completed an out-of-state teacher preparation program but who has not yet successfully completed the Praxis I series or Core Academic Skills for Educators basic skills test and the state's Praxis II competency tests in the content area.

West Virginia summative Assessment Achievement Levels – Level 1, Level 2, Level 3, and Level 4.

For Grade 5

Level 1 – The student has not met the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in mathematics needed for likely success in future coursework.

Level 2 – The student has nearly met the achievement standard and may require further development to demonstrate the knowledge and skills in mathematics needed for likely success in future coursework.

Level 3 – The student has met the achievement standard and demonstrates progress toward mastery of the knowledge and skills in [content area] needed for likely success in future coursework.

Level 4 – The student has exceeded the achievement standard and demonstrates advanced progress toward mastery of the knowledge and skills in [content area] needed for likely success in future coursework.

For Grade 11

Level 1 – The student has not met the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after high school.

Level 2 – The student has nearly met the achievement standard and may require further development to demonstrate the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after high school.

Level 3 – The student has met the achievement standard and demonstrates progress toward mastery of the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after completing high school coursework.

Level 4 – The student has exceeded the achievement standard and demonstrates the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after high school.

Limitations

A teacher's Highly-Qualified designation may vary based on the regulations in place at the time the designation was received making it difficult to differentiate how different teachers with the same designation achieved such Highly Qualified designation.

Candidates may have additional out-of-state years of teaching experience not captured by the data management system utilized throughout West Virginia. This out of state experience may potentially place these individuals in an incorrect bracket regarding years of experience.

Results of the West Virginia standardized summative assessments in English and mathematics may be affected by certain social and economic circumstances of the students. There are teacher and student characteristics that could affect student achievement; however, this study's scope was limited to only certification status, teacher experience, highly qualified status, and preparation pathway.

Delimitations

The study included a selection of a statistically significant sample of WV teachers prepared by traditional routes and alternative routes. In addition, it included West Virginia summative assessment performance levels from all 5^{th} and 11^{th} grade students taught by the selected sample of teachers for the 2015-16 academic year.

Research was conducted within the context of a specific area, West Virginia, and a specific set of teacher quality variables that included teacher preparation, licensure, Highly Qualified status, years of experience, and grade levels. Within this study, learning and growth are referenced as the results of standardized test scores in mathematics and English language arts.

Rationale

Depending on the results, the outcome of this research may address specific teacher certification needs in West Virginia and help inform local policy as well as explore potential funding implications. As West Virginia currently has over 700 teacher vacancies, the findings would assist state and local education agencies as well as teacher preparation programs to tailor traditional and alternative preparation programs to address best practices identified that may have a significant impact on student achievement. In order for legislators and local and state education agencies to implement effective strategies and to scrutinize the characteristics that may produce

the necessary growth in student learning, it is vital to identify the factors which influence teacher effectiveness and which have the most potential positive impact on such learning.

The information obtained from the results may help determine if the teacher quality variables in question do directly influence student learning and growth as evidenced by the related standardized test results in mathematics and English language arts. These results may show if the standardized achievement results for students in West Virginia are significantly impacted by how their teachers are prepared, by what type of licensure their teachers hold, and by their Highly Qualified designation and how the teachers achieved such designation. No matter the results, all involved at the state or local level would be better able to revise their hiring policies, teacher assignments to specific vacancies, and tailor further professional development for their teachers. Additionally, the findings may help legislators guide state statutes regarding educator preparation and teacher evaluation and compensation.

If the results for the effect of teacher quality are significant, then the state could implement those qualities as requirements for all new teachers and teacher education programs would need to make revisions accordingly. Preparation programs could also provide related professional development for in-service teachers not meeting the designation. If such results are not significant or inconclusive, then the state could reconsider these designations and potentially develop a new set of characteristics.

If the results show significance for either type of teacher preparation program—alternative or traditional—then that pathway can continue to be developed and refined, particularly for addressing critical shortage areas. If these pathways are inconclusive, then the state could deemphasize alternative preparation or design newer approaches and may assist financially given the existing structures of traditional programs.

If teacher experience is a significant effect, favoring the experienced teacher, then the state could consider enhancing teaching or mentoring models where experienced teachers are paired with new teachers in either a professional development model or a tiered compensatory teaching model. If not significant, then the current model can be continued with the beginning teacher induction program. Additionally, if there is no significance on student test scores, then the state might consider alternative assessments that are performance based, with repeated measures, rather than a one-time measure that significantly affects policy.

If significance is found for one content test area rather than the other, and for one grade level than the other, then administrators and curriculum developers could study such results in depth and make relevant revisions to either the standards and/or design school-based curriculum strategies for enhancing the relevant content. Also, the state could begin to examine the seriousness with which, for example, high school and middle school students approach test taking to determine if optimal scores are being obtained. Often these scores are related to school compliance measures and the state and the public should be assured that these kinds of decisions are being made with reliable and valid data.

If type of licensure is significant and favors, e.g., the professionally certified teacher, then the state can be confident about the requirements of this model for certifying teachers and about structuring the existing requirements for temporarily certifying teachers, but perhaps strengthening these and/or creating a more closely supervised context for these teachers.

Finally, the results of this study might help contextualize the ranking West Virginia receives in national publications such as the National Council on Teacher Quality (NCTQ) and the National Assessment of Educational Progress (NAEP) in order to better position policy

makers and state education agencies to seek both private and public competitive funding opportunities.

Summary

In closing, current research shows that issues of teacher quality and its effects on student achievement are in flux nationally and being promulgated heavily by enactments of the federal government and its various policies as well as by private entities who provide public rankings of states and teacher preparation programs. Literature supports the need for further research regarding student academic achievement and teacher qualities. These qualities need to be very specifically defined and filtered to account for variables that may interfere with usable findings.

Finally, such information and data, no matter the results, are important contributions to a state and a national database on the relationship between teacher qualities and student achievement.

CHAPTER 2

REVIEW OF LITERATURE

This study will expand upon existing research as it relates to academic achievement and its relationship to teacher licensure, preparation pathway, years of experience, and highly qualified designation. It will examine the variables separately and then how these connect or overlap. Furthermore, it will identify the factors that have shown to have an impact on academic achievement as well as those that did not significantly impact it. Overall, this review is organized by four areas: student academic performance and teacher preparation route, student academic performance and teacher years of experience, student academic performance and state teacher licensure, and student academic performance and highly qualified designation of teacher.

The studies selected focused on a number of variables related to student learning and the characteristics of their teachers. These studies included the quality of the teacher, the type of teacher preparation program and its resulting type of licensure. Additionally the effect of teacher experience on student achievement was reviewed, including the effects on students at the middle and high school levels.

Much of the existing research on teacher quality and preparation pathways has been aimed at particular school districts or contexts in rural and urban areas, rather than being statewide or national investigations. The majority of the studies were quantitative investigations although several had qualitative components as well. The studies selected were conducted between 1999-2015, with the majority completed between 2000-2009.

Student Academic Performance and Teacher Preparation Route

When researching what makes an effective teacher, many point to the initial preparation of the teacher as having the most impact on their effectiveness (Darling-Hammond, 2000;

Peterson and Nadler, 2009). Teachers are often said to be the most critical factor when it comes to student learning. In addition, there are other background characteristics such as ethnicity, socio-economic status, and parental influence and background that have a large impact on learning. However, available research has not yet provided a direct link between variance in student achievement and any particular background characteristic. The major portion of variance in achievement has been directly linked to the impact of the teacher (Sawchuk, 2011).

Goldhaber and Brewer (1999) discuss the fact that not all subject areas are equal when it comes to the degree of influence the teacher has when it comes to how students score on achievement assessments. The authors further found that all the combined variables affecting achievement resulted in 21% of the variation in mathematics achievement. Additionally he noted that about 8.5% of such variance was directly attributed to the teacher's influence. It is important to note that there are other areas which have not been studied as much that also would have an impact on students such as administrative involvement and influence of the principal and other administrators. Also, how district initiatives are interpreted and implemented by the teachers can have a significant impact on achievement (Rothstein, 2010).

Teacher preparation and accomplishments are also viewed as very influential variables regarding student academic performance. The National Commission on Teaching & America's Future (2016) reports that teachers who complete rigorous programs including performance-based processes such as the National Board Certification help students make gains comparable to as much as 2 months of learning. Academic achievement may also be impacted by current teacher shortages and teacher preparation enrollment decline across the country. As indicated in the Title II reports by the U.S. Department of Education (2015a), during 2013-14 teacher preparation completion rates dropped by over 123,000 across the United States. Such shortages

create situations where students are potentially taught by a variety of teachers within the same year. These teaching scenarios could range from having a short-term or long-term substitute or someone pursuing alternative certification being the teacher of record. When a more permanent teacher is hired into one of those classrooms, it would be difficult to determine the level of success of any of the individual teachers who taught in that classroom.

Mentorship during the pre-service and in-service period needs to be considered. Some of the differences between preparation routes are the type and amount of mentorship teacher candidates receive. Teachers who have more mentorship and/or induction perform more effectively (Ingersoll & Strong 2015; Snyder & Bristol, 2015). Some alternative pathways have required mentorship embedded throughout the program as candidates complete their programs while on the job. Hence there is the need to separate the different pathways of teacher preparation when looking at impact on student achievement.

When considering programmatic levels, Clofelter, Ladd, and Vigdor (2010), studied the impact teacher credentials had on student achievement at the high school level. One of the most significant findings showed that the type of credential held by the teacher affected how students achieved. The impact was significant enough to result in state policy changes. At the elementary level across the country, in the areas of mathematics and reading, students who were taught by certified teachers outperformed their peers who were taught by teachers who were not certified (Riordan, 2009).

It is evident that much of this effort to understand and identify what best promotes academic achievement reaches beyond the United States. Many studies have been conducted across the world trying to address these same questions. Recently, as indicated by the Program for International Student Assessment (PISA, 2015), well-prepared teachers and how they are

prepared have a significant impact on achievement. Countries with high performing schools and students have well-prepared and highly-skilled teachers (Schleicher, 2013). When reviewing the literature nationally and internationally, it is evident that additional research is needed that tie together many of these variables that impact student achievement. As Ingersoll, Merrill, & May (2014) concluded, teacher education preparation is directly and significantly related to how well students achieve.

Student Academic Performance and Teacher's Years of Experience

Experience is often viewed as desirable and at times a required element under most circumstances. Experience is required for most jobs, trades, and even volunteering opportunities. Education and teaching are also held to such beliefs. Experience factors into such aspects of the workforce including salary, tenure, and benefits. The belief is that experience improves effectiveness, which delivers better results (Rice, 2010). Existing research indicates that experience alone and amount of experience are not the only or greatest determining factors of effectiveness and quality. To determine impact significance one must look in greater detail and at additional factors that influence the outcome as well (Clotfelter, Ladd, & Vigdor 2007). This study will look in greater detail how and if such compounding factors have an impact on the academic performance of students.

Many have examined additional factors that potentially influence student achievement. Budding & Zamarro (2009) reported that how teachers score on licensure tests and advanced degrees has no impact on achievement; however, teacher experience does. Often those with greater experience also hold advanced degrees but the researchers report that degrees alone have no statistical significance. When looking at some of these variables differently, Clotfelter, Ladd,

& Vigdor (2007a) concluded that teacher licensure scores have a significant impact on mathematics achievement but teacher experience is a constant by having impact across all areas.

Rice (2003) describes experience being a key element but further indicates that greater significance occurs at the secondary level rather than at the elementary level. Across the country these results differ. In Ohio for example, according to Carr (2006), teacher experience was not statistically significant for impacting student achievement. Huang & Moon (2009) and Harris & Sass (2007, 2013) on the other hand, found that teacher experience only had a significant impact on particular grade levels but not across different grade levels. Ladd & Sorensen (2014) reviewed teachers' years of experience across a longitudinal study in North Carolina and concluded that teacher experience had a clear impact on how students scored on tests but also looked and identified other positive effects such as improvement in student behavior and a significant reduction in absenteeism.

One other factor often cited and described in the available literature is the relationship between instructional approaches and the teacher's years of experience and the impact on teacher effectiveness (Smith, Lee, & Newmann, 2001). Many of the conclusions provided in the literature, including longitudinal studies, have focused on a particular grade level or content area. Wiswall (2013) found that 5th grade classroom teachers do not produce better results in student achievement after the initial first few years of teaching regardless of their years of experience. Such findings demonstrate how narrow the focus traditionally is when looking at just individual or few variables. It is imperative to consider the cumulative effects the aforementioned variables have on student achievement in their different permutations.

Student Academic Performance and Type of State Teaching Licensure

Darling-Hammond (2000), reports that teacher certification and licensure have a direct impact on increased student achievement. She also found that the NAEP scores in some areas such as mathematics are lower for students taught by non- fully certified teachers. Several studies also point to the importance of the type of teacher credentialing and how those impact achievement. Darling-Hammond (2007) found that not all teacher credentials have equal impact on achievement. She describes certain credentials such as alternative certification as having a negative impact on student achievement. This negative impact may be partially attributed to the turnover of some of the alternatively certified teachers. Perhaps if the alternatively certified teacher stayed longer, over time, they would have greater impact on the academic achievement of students. Others (Kane, Rockoff, & Staiger, 2008, Goldhaber & Brewer, 2000) argue that there is no significant difference on achievement levels when linked to teacher certification type.

Licensure is one of the key factors and primary requirement states use throughout the country to uphold and justify the quality of their teachers. States differ when it comes to licensure requirements, programs leading to licensure, number of hours required and the types of licensure issued (Goldhaber & Brewer, 2000). Wenglinsky (2000) analyzed data from NAEP and reported that teachers who had a major or minor in the content area they taught had a greater impact on achievement rather than licensure. Some argue that teacher quality is a significant factor but it is not related to the type of licensure teachers hold (Koedel & Betts, 2007). In a study by Aaronson, Barrow, & Sander (2007) the authors found that the teacher has a significant impact on student achievement but individual characteristics including certificate type do not.

Many studies do not separate all types of credentials and advanced credentials equally. Cowan and Goldhaber (2015) indicate that when looking at individual characteristics, National

Board Certification has a greater impact on student achievement when compared to any other characteristic. The authors also found that teachers who perform better on the National Board Certification assessments have greater effectiveness ratings. Based on these collective findings, it appears that further clarification when studying certification type needs to be provided. Such clarification should include more details on the actual certification type either initial, advanced, or nationally such as the National Board for Professional Teaching Standards certification. As stated in the purpose of this study, including specific state teacher certification in combination with other teacher characteristics may provide greater data on its impact on student achievement.

Student Academic Performance and Highly Qualified Designation of Teachers

Since its inception with the No Child Left Behind Act in 2001, the debate over what is a highly qualified teacher continues. This piece of legislation set out to staff a highly qualified teacher in every classroom by 2006. Inclusion of such a requirement in the legislation implies the federal government is confident that there is a strong correlation between highly qualified and teacher quality (Holloway, 2007). According to Rothman (2009), even eight years after its inception, even though most teachers across the country have met the HQ definition, there is no significant evidence that the quality of teachers changed. The discrepancies over such goals and claims are many. HQ status may vary from state to state. The legislation provided a large range of flexibility in how states defined the Highly Qualified (HQ) status of a teacher. The provided flexibility may vary from a prescribed number of academic hours in a particular content area to designations given by the school principal based on classroom observations (U.S. Department of Education, 2004; Holloway, 2007).

A primary criterion of the Highly Qualified designation is proper certification in the content area taught. Goe (2007) and Betts and Frost (2000) indicate that teachers licensed in the

content area they teach show greater impact on how students achieve. Furthermore, Nyankori (2005) and Cronigner, Rice, Rathun, & Nishio (2003) argue that teacher certification has an indirect impact on achievement; however, the authors focused more on the impact of the combined characteristics of teachers for the entire school and the related impact on achievement. It is evident that many of the studies have focused on how the combination of several of these variables impact achievement. When it comes to the HQ status of teachers, most available research analyzes separate variables that collectively result in an HQ designation but not at the designation as a whole.

As indicated in this review, there is a need for further research in this specific area as it connects to HQ status. Specifically, the area of focus should be on the impact that designation of HQ as a whole has on achievement. Several contradicting bodies of research including Darling-Hammond (2000a), Darling-Hammond and Sykes (2003) and Ryan (2004), indicate that the Highly Qualified designation as a whole may not be an accurate measure of effectiveness and quality. These authors concluded that more careful attention to individual teacher characteristics would be more beneficial to policymakers and school systems when determining teacher effectiveness and any potential impact it may have on student achievement. It is the purpose of this study to analyze in greater detail how these unique variables impact student achievement.

Summary

Academic achievement is one of the most difficult outcomes to measure because it has a very large number of variables affecting it. Several of the education constituencies including federal and state government, and local school districts believe that measuring academic achievement via standardized testing is the most reliable way to provide data regarding the efficacy of education (Gawthrop, 2014). Sanders (1998) stated that the most important factor

impacting academic achievement and growth is teacher effectiveness. Such findings are found throughout the now reauthorized Every Student Succeeds Act (U.S. Department of Education, 2015) and its previous version the NCLB Act (U.S. Department of Education, 2014).

Considerable emphasis has been placed on these variables trying to find the right combination in order to impact student achievement as well as to better prepare teachers. Academic achievement is also used as a measure of school district success. Consequently, the variables related to such achievement are under continuous scrutiny (Brinkman, 2014). The characteristics of educators are the focus of a large body of research to determine the impact on student achievement (Sanders & Rivers, 1996). Therefore, there is great need to further explore the literature relevant to how these variables impact student achievement.

A consistent and clear message across all research is that a common meaning of teacher quality has not been reached but teacher quality is without a doubt the most cited factor impacting student achievement (Goe, 2007). The preponderance of the available research in this area often only addresses individual variables primarily across one of the programmatic levels: elementary, middle, or secondary. Much of the research focuses on particular school districts or particular characteristics of an area such as urban or rural. There is a gap in how the same variables and factors impact achievement across these programmatic levels considering variables such as teacher licensure, preparation pathway, years of experience, and highly qualified designation.

CHAPTER 3

METHODS

This chapter will initially describe the purpose of the study and its major treatment and outcome variables. Additionally, it outlines its major components including design, population and subject selection, research questions, major procedures, and data analysis.

Purpose

The purpose of the study is to determine the effect of traditional and alternative teacher preparation, years of service, type of licensure held and Teacher quality on West Virginia English language arts and mathematics summative assessment scaled scores and performance levels among students in grades 5 and 11. Specifically, this research will analyze the existing assessment data in West Virginia's 55 counties, regarding teacher experience, the types of license held including professional licenses and alternative licenses, and the teacher's Highly Qualified (HQ) designation. What might be the effect, if any, on student achievement that can be attributed to the teacher quality variables noted previously?

Design

This study collected licensure, employment, and assessment data to examine the level of impact on grade level standardized tests in mathematics and English language arts moderated by the kind of teacher preparation, years of teacher service, type of licensure held and highly qualified status. The major outcome variables are English language arts and mathematics scaled scores and related categorical rankings on the summative, year-end assessment in West Virginia.

Population

The study includes all mathematics and English teachers who were currently employed in a WV public school in all 55 counties during academic year 2015-2016 in grades 5 and 11. These subjects were obtained from an encrypted statewide database at the West Virginia Department of Education (WVDE). The data were received and approved through a formal data request submitted to the WVDE.

There are two sub-populations in the design: all 3,589 mathematics and English classroom teachers in grade levels 5 and 11, in West Virginia's 55 counties during the 2015-16 school year, and their respective numbers of English/language arts and mathematics students in grades 5 and 11 for a total of 34,528. However, 155 teachers were omitted because their preparation pathway could not be clearly identified as traditional or alternative. In all, there were 3,434 included in the population.

It is recognized that within the sample there is an overlap in the variables and factors of the study. For example, teacher experience encompasses all participants in the study no matter the variable (certified, noncertified, highly qualified, non-highly qualified, alternatively prepared and traditionally prepared).

Research Questions and Data Analysis

The following research questions overarch and guide this study. Data analysis will include a combination of descriptive and inferential statistical techniques. Each of these are noted below and aligned with the study's respective research questions.

1. What are the differences between the West Virginia summative assessment scaled scores in English language arts and mathematics among West Virginia students in grades 5 and 11, taught by teachers trained in traditional and alternative teacher preparation programs?

2. What are the differences in categorical rankings (levels 1, 2, 3, and 4) on the West Virginia summative assessment in English language arts and mathematics among West Virginia students in grades 5 and 11 taught by teachers trained in traditional and alternative teacher preparation programs?

3. What is the effect of teacher experience on West Virginia assessment summative scaled scores in English language arts and mathematics among West Virginia students in grades 5 and 11?

4. What is the effect of teacher experience on West Virginia summative assessment categorical rankings (levels 1, 2, 3, and 4) for mathematics and English language arts, grades 5 and 11?

5. What is the effect of the type of teacher licensure (certified/noncertified) on West Virginia summative assessment scaled scores in English and mathematics among West Virginia students in grade 5 And 11?

6. What is the effect of type of licensure (certified/noncertified) on categorical rankings in math and English, grades 5 and 11?

7. What is the effect on math and English scaled scores for 5^{th} grade math and English teachers with highly qualified status or non-highly qualified status?

8. What is the effect of highly and non-highly qualified teacher status on categorical rankings in English and mathematics for grades 5 and 11?

Procedures

Data Request

The data was requested from the West Virginia Department of Education by following a data request protocol established by the agency. A written request was submitted to the Office of Data Governance through the Zoom WV portal. The request included the description of the data

including year, grades, type of licensure held, years of experience, and highly qualified status. The data was provided at the aggregate level for each grade band requested as an Excel spreadsheet (assessment_data_suppressed_Hagerman.xls) suppressed and de-identified to remove all identifiable information. The approval copy is available in Appendix A.

Data Collection and Match

WVDE staff matched the category of teachers requested to the students they taught. After the match was completed, all identifiable information was removed and data was grouped by categories and variables. All data including teacher licensure, years of experience, and Highly Qualified designation were provided at the aggregate level so no identifiable information for both teachers and students was provided. As a result, the database did not require any secured features or permissions to be accessed. However, the data was kept in a password protected external drive known only to the researcher.

Data Import

The data was imported to SPSS, Version 24 from the Excel spreadsheet for analysis. Other than having to name the variables at the Data Screen, the input into SPSS is ready to perform the various kinds of analyses needed. Before and after the analysis process the data were stored in a password protected external drive as well as at the Output database on SPSS. All data were analyzed in the aggregate form and no names or identifiable information will be available.

IRB Protocol

The research prospectus was submitted to the candidate's doctoral committee and approved by the committee. Subsequently, it was submitted to the Marshall University Institutional Review Board (IRB) for approval. The IRB Research (Protocol) Application, Form #2 (Social/Behavioral) was submitted to the Marshall University Institutional Review Board for

review and approval. Following these reviews, the data was analyzed in SPSS, Version 24, for the appropriate statistical models.

Data Analysis

Based on the existing research questions for the investigation, analysis will be obtained by a combination of descriptive and inferential statistical methods. Descriptively, data included aggregate mean score data for the various variables and groupings, with related standard deviations and standard error scores. Inferential analysis included one-way analysis of variance, t-tests for independent samples and nonparametric models including Kruskal-Wallis and Chi-Square tests of independence. These data will also be supported with various graphic details and representations.

In each case, the test of significance will be set apriori at a p-level < or equal to .05 and related effect size measures will be obtained for results that are statistically significant. English language arts and mathematics scaled scores will initially be analyzed descriptively by obtaining means of the scaled scores, standard deviations, and measures for skewness and normality. Data will be inferentially analyzed using a combination of t-tests for independent samples and analysis of variance. Related categorical rankings will be analyzed using a combination of nonparametric tests: Chi-Square Goodness of Fit and Kruskal-Wallis.

Summary

The central purpose of the investigation is to determine how student achievement in mathematics and English/Language Arts might be distinguished by a set of teacher characteristics. These include how the teachers were academically trained, vis a` vis traditional teacher preparation programs or alternative pathways to certification. Additionally, does the experience of the teacher interplay with one’s type of academic preparation and level of formal

teacher certification? Notwithstanding these purposes, it was of prime importance to know something about the quality of the teacher, with the belief by many that this variable is strongly correlated to student achievement. While the several research questions posed will be examined as separate effects, the combined effects of the variables as a whole may be more revealing. The significance of the results, once known, may potentially shed some light on the connection of WV teachers and their professional profiles to their students and their achievement status.

The entire data set associated with the respective variables will be obtained from a large, fully protected, encrypted database from a state agency. Consequently, the investigation will avoid some of the pitfalls of real time sampling: bias selection, mortality, and inadequate sample sizes, and lacking compliance for security and confidentiality.

CHAPTER 4

DATA ANALYSIS AND RESULTS

Purpose

The purpose of the study is to determine the effect on English language arts and mathematics scaled scores and related categorical performance levels on the West Virginia summative assessment among students in grades 5 and 11, moderated by traditional and alternative teacher preparation, years of experience, type of licensure held, and Highly Qualified designations. Specifically, this research will analyze the existing assessment data in West Virginia's 55 counties to determine what might be the effect, if any, on academic achievement that can be attributed to these teacher variables.

Data

The data for the study included scaled test scores and categorical ranked median scores for mathematics and English language arts among a statewide sample of students in grades 5 and 11. Test score data were obtained from the WVDE, housed at the state level in the West Virginia Zoom Data Warehouse. The data for the warehouse were collected through the WV Education Information System (WVEIS), a secure database and system within the WVDE servers.

Population/Sample

The population for this study included 3,589 5^{th} and 11^{th} grade English/language arts and mathematics teachers in the state of WV. These are WV teachers assigned as a teacher of record in the master schedule for all schools in WV that contain a 5^{th} grade and an 11^{th} grade classroom. The population was comprised of 1,777 5^{th} grade teachers that included 868 math teachers and 909 English/language arts teachers and 1,657 11^{th} grade teachers that included 789 math teachers and 868 English/language arts teachers. These were pulled from the master schedules for every

school meeting the criteria (5th and 11th grades with math and English/language arts teachers). Because their preparation route was unable to be determined, 155 teachers were omitted from the sample.

The student population for this study included every student assigned to every 5th and 11th grade teacher in the above sample, who took the assessment in the 2015-2016 academic year. The student sample included 17,546 students in mathematics for grade 5, and 17,495 students in English/language arts 5th grade, 15,502 mathematics 11th grade students, and 16,116 English/Language arts students for grade 11. Of note, students within the same grade band (5 or 11) may appear in both the math and English language arts groupings.

Because several of the groupings for each variable of the study varied in size and were disproportional, random sampling was obtained for each of the variables shown in Table 1.

Table 1

Sample Sizes for Variable Factors Across Grade Levels and Content Areas

	ELA 5th	ELA 11th	Mathematics 5th	Mathematics 11th
Alternatively Certified	416	1,355	263	1,378
Traditionally Certified	364	1,446	348	1,418
	ELA 5th	ELA 11th	Mathematics 5th	Mathematics 11th
Certified	364	3,218	348	642
Non-Certified	287	3,005	242	797
	ELA 5th Grade	ELA 11th Grade	Mathematics 5th Grade	Mathematics 11th Grade
Highly Qualified	120	814	115	487
Non-Highly Qualified	108	737	95	460

Research Questions and Data Analysis

Research Question #1:

"What are the differences between the West Virginia summative assessment scaled scores in English language arts and mathematics among West Virginia students in grades 5 and 11, taught by teachers trained in traditional and alternative teacher preparation programs?"

This question addressed differences in scaled scores on the West Virginia summative assessment in mathematics and English among 5th and 11th graders taught by teachers trained in traditional compared to alternative teacher preparation programs. In each case an independent samples t-test was obtained to statistically test the significance of these variables. Table 2 shows the group statistics test data for math scaled scores among 5th graders moderated by the teacher preparation variable.

Table 2

Group Statistics Data for Alternatively or Traditionally Certified 5th Grade Math Teachers

Alt/Trad		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	0 = Trad	348	2489.71	80.446	4.312
	1 = Alt	263	2447.42	99.820	6.155

Numerical data in the group statistics table 2 shows a difference of approximately 42 scaled score points favoring traditionally prepared, 5th grade Math teachers. Additionally, an effective difference in the variability of the standard deviation scores occurred for alternatively prepared teachers sample sizes were effectively equitable.

Table 3

Independent Samples T-Test for 5th Grade Math Teachers, Alternatively or Traditionally Certified

		Levene's Test for Equality of Variances		T-Test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
STUASM	Equal variances assumed	23.410	.000	5.797	609	.000	42.294	7.296	27.966	56.623
	Equal variances not assumed			5.628	492.701	.000	42.294	7.516	27.528	57.061

The data in Table 3 indicates a significant statistical effect for traditionally prepared 5th grade mathematics teachers compared to their alternatively prepared peers, with a mean score difference of approximately 42 points, and a p level of .000 or <.0005. An effect size measure of .05218 indicated approximately 5% of the variability was accounted for by the predictor variable (Cohen, 1992). In effect, students taught by traditionally prepared 5th grade mathematics teachers scored significantly greater on the WV summative content exam than did their peers taught by alternatively prepared mathematics teachers.

Table 4

Group Statistics Data for Alternatively and Traditionally Prepared 11th Grade Math Teachers

Group Statistics

Alt/Trad		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	0 = Trad	1418	2538.99	119.297	3.168
	1 = Alt	1378	2519.33	120.783	3.254

The data in Table 4 resulted in a mean score difference of approximately 20 scaled score points favoring 11th grade traditionally prepared teachers. Variability was stable with small differences between standard deviation and standard error values. Sample sizes for the groupings were essentially equitable.

Table 5

Independent Samples T-Test for Alternatively and Traditionally Prepared 11th Grade Math Teachers

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	18.197	.000	5.582	778	.000	38.387	6.877	24.887	51.887
Equal variances not assumed			5.646	776.764	.000	38.387	6.799	25.041	51.733

The data in Table 5 indicates a statistically significant effect for traditionally prepared 11th grade mathematics teachers compared to their alternatively prepared peers, with a p level of .000 or <.0005. An effect size measure of .0411 indicated approximately 4% of the variability was accounted for by the predictor variable (Cohen, 1992).

Table 6

Group Statistics Data for Alternatively and Traditionally Prepared 5th Grade English Teachers

Alt/Trad		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	0 = Trad	364	2498.54	86.690	4.544
	1 = Alt	416	2460.15	103.147	5.057

The data in Table 6 resulted in a mean score difference of approximately 38 scaled score points for students who were taught by traditionally prepared 5th grade English teachers compared to peers who were taught by alternatively prepared English teachers. There were greater variability values associated with alternatively prepared teachers, or a slightly greater standard error around the mean and the same for standard deviations. The sample sizes for the groupings were effectively equivalent.

Table 7

Independent Samples T-Test for Alternatively and Traditionally Prepared 5th Grade English Teachers

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
ENGLISH STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	18.197	.000	5.582	778	.000	38.387	6.877	24.887	51.887
Equal variances not assumed			5.646	776.764	.000	38.387	6.799	25.041	51.733

A significant statistical effect was found for traditionally prepared 5th grade English teachers compared to their alternatively prepared peers, with a mean score difference of approximately 38 points, and a p level of .000 or <.0005. An effect size measure of .082 indicated approximately 8% of the variability was accounted for by the predictor variable.

Table 8

Group Statistics for Teacher Preparation Variable, 11th Grade English Teachers

Alt/Trad	N		Mean	Std. Deviation	Std. Error Mean
STUASMSCO 0 = Trad	1446		2576.43	110.038	2.894
1 = Alt	1355		2560.50	113.842	3.093

These data show a mean score difference of approximately 16 scaled score points favoring students who were taught by traditionally prepared 11th grade English teachers compared to their alternatively prepared peers. Sample sizes for the groupings were essentially equivalent for proportional data analysis.

Table 9

Independent Samples T-Test for Teacher Preparation Variable, 11th Grade English Teachers

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	2.535	.111	3.766	2799	.000	15.932	4.231	7.637	24.228
Equal variances not assumed			3.762	2771.855	.000	15.932	4.235	7.628	24.237

A statistically significant finding occurred for traditionally prepared 11th grade English teachers with a mean score difference of approximately 16 scaled score points, and a p level of

.000 or <.0005. An effect size measure of .091indicated approximately 9% of the variability was accounted for by the predictor variable (Cohen, 1992).

Summary

Analysis of the data for both grade bands, 5th grade and 11th, indicated that there is a statistically significant difference favoring teachers prepared in a traditional manner when compared to those prepared through an alternative pathway. When looking at the individual grade bands, 5th grade showed the biggest difference. Those teaching 5th grade English showed a difference of 38 scaled score points for traditionally prepared and those teaching 5th grade mathematics showed a difference of 42 scaled points. Teachers of 11th grade mathematics showed the largest difference by 20 scaled score points for those traditionally prepared and those teaching 11th grade English showed a difference of 16 scaled score points.

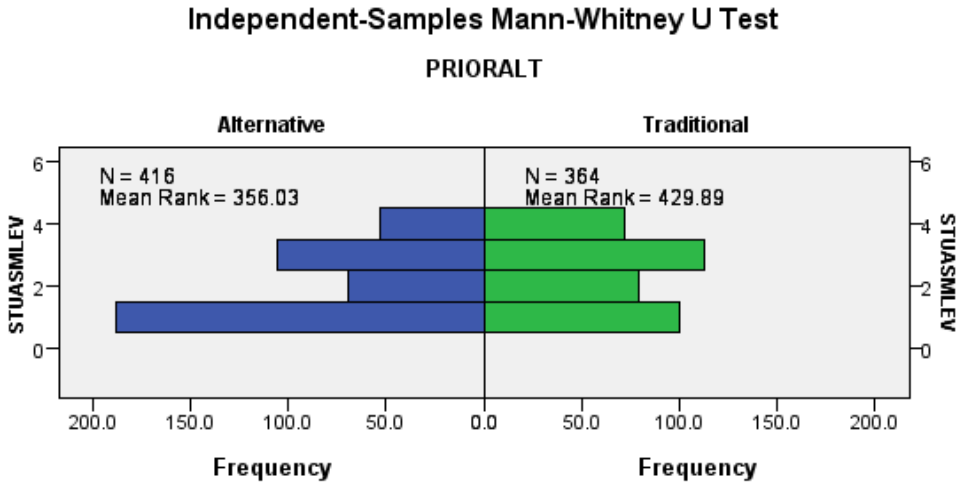
In effect, the overall results indicated that 5th and 11th grade WV students taught by traditionally prepared teachers scored significantly greater on WV summative content exams than did their peers taught by alternatively prepared teachers, notwithstanding the content area.

Research Question #2:

"*What are the differences in categorical rankings (levels 1, 2, 3, and 4) on the West Virginia summative assessment in English language arts and mathematics among West Virginia students in grades 5 and 11 taught by teachers trained in traditional and alternative teacher preparation programs*?"

In each case, these outcomes were assessed using a two factor, nonparametric statistical model (Mann-Whitney U) designed to test the significance of ranked median data. Figure 1 shows the group statistics data test for English scaled scores among 5th graders moderated by the teacher preparation variable.

English 5



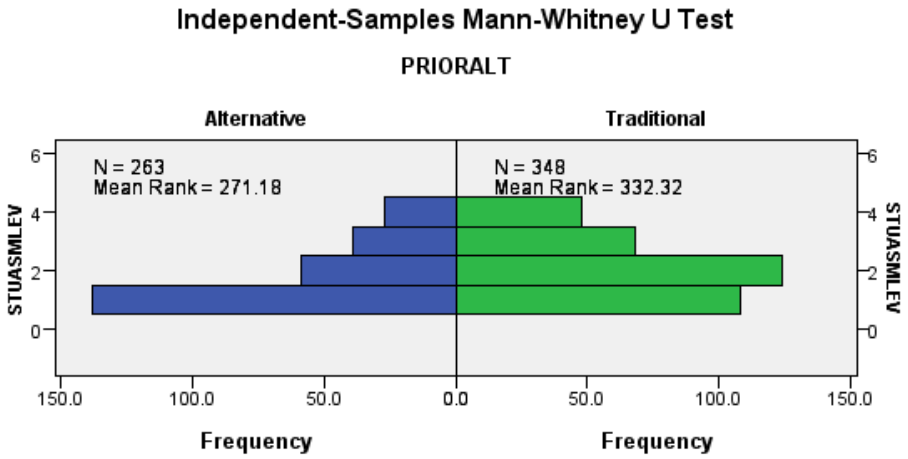
Total N	780
Mann-Whitney U	61,372.500
Wilcoxon W	148,108.500
Test Statistic	61,372.500
Standard Error	3,005.510
Standardized Test Statistic	-4.771
Asymptotic Sig. (2-sided test)	.000

Figure 1. Visual and numerical data comparing categorical rankings and performance levels among alternatively and traditionally prepared 5th grade English teachers

These data show that 5th grade traditionally prepared English teachers ranked significantly greater across the four performance levels, with a difference of 74 ranked scores compared to alternatively prepared teachers. Likewise, the test of significance table confirmed statistical significance with a z approximation test score of 4.771 and a test value of p .000 or <.0005. In effect there is a statistically significant difference, far beyond chance, in ranked scores

favoring traditionally prepared teachers. It is interesting to note that the frequencies for level 1 in the ranks figure (standard not met) are nearly doubled for alternatively prepared compared to traditionally prepared. An effect size measure of .1025 indicated that approximately 10% of the total variability was accounted for by the predictor variable.

Math 5



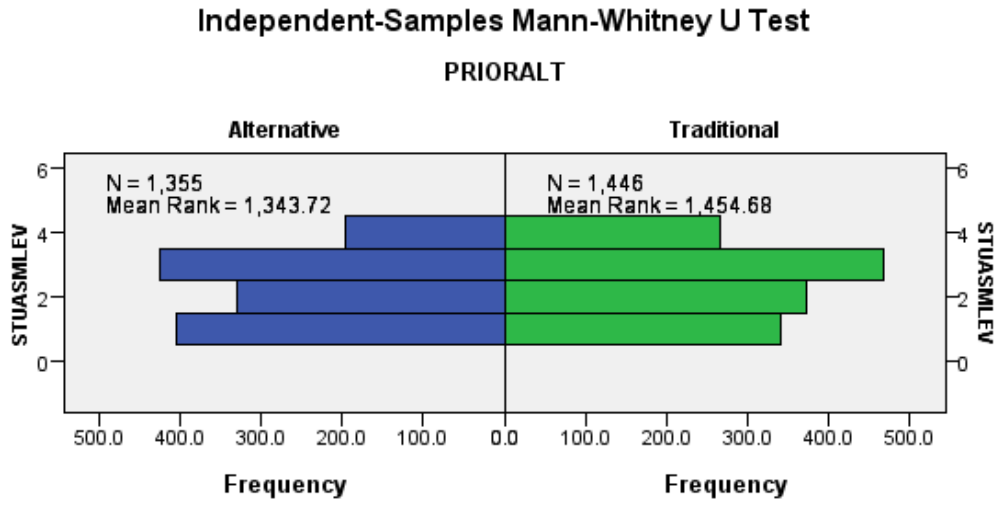
Total N	611
Mann-Whitney U	36,604.000
Wilcoxon W	71,320.000
Test Statistic	36,604.000
Standard Error	2,050.362
Standardized Test Statistic	-4.467
Asymptotic Sig. (2-sided test)	.000

Figure 2. Visual and numerical data comparing categorical rankings and performance levels among alternatively and traditionally prepared 5th grade math teachers

Figure 2 shows that traditionally prepared 5th grade math teachers outranked their alternatively prepared peers by 85 ranked scores which resulted in a significant difference (z,

4.467, p .000 (<.0005). An effect size measure of 12.2 accounted for about 12% of the total variability. It is observable that the combined frequencies for performance level 1 in the ranks table (standard not met) and level 2 (nearly met) were greater for traditionally prepared teachers, notwithstanding their overall significance.

English 11

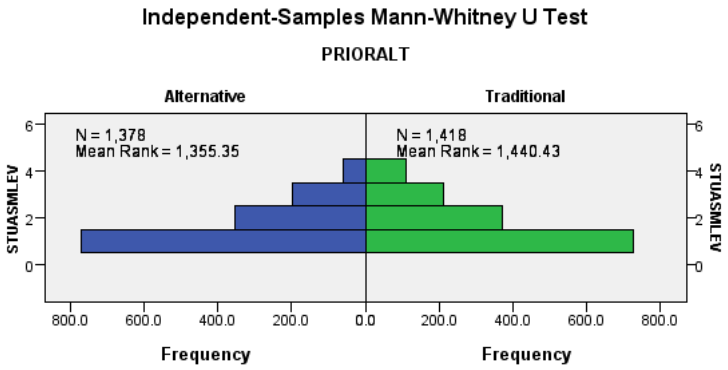


Total N	2,801
Mann-Whitney U	902,046.000
Wilcoxon W	1,820,736.000
Test Statistic	902,046.000
Standard Error	20,612.139
Standardized Test Statistic	-3.766
Asymptotic Sig. (2-sided test)	.000

Figure 3. Visual and numerical data comparing categorical rankings and performance levels among alternatively and traditionally prepared 11th grade English teachers

The data in Figure 3 showed that traditionally prepared 11th grade English teachers outranked their alternatively prepared peers by 111 ranked scores which resulted in a statistically significant difference (z, 3.766, p.000 (< .0005). An effect size measure of 0.075 accounted for about 8% of the total variability. Although a significant effect occurred overall, it is observable that, for both groupings, the combined frequencies for level 1 (standard not met) added up to nearly 760 or about 27% of the population of 11th graders.

Math 11



Total N	2,796
Mann-Whitney U	917,539.500
Wilcoxon W	1,867,670.500
Test Statistic	917,539.500
Standard Error	19,398.497
Standardized Test Statistic	-3.065
Asymptotic Sig. (2-sided test)	.002

Figure 4. Visual and numerical data comparing categorical rankings and related student performance levels among alternatively and traditionally prepared 11th grade math teachers

Similarly, Figure 4 shows traditionally prepared 11th grade mathematics teachers outranked their alternatively prepared peers by 43 ranked scores which resulted in a significant

difference (z, 3.056, p .002 (<.0005)). An effect size measure of 0.057 accounted for about 6% of the total variability. In this case, it is interesting to note that the frequencies in performance levels 1 (standard not met); 2 (almost met) and 3 (standard met) were nearly identical for both groupings.

In addition to the test of significance for the teacher preparation variable, continuous field data were examined in regard to the frequencies of scaled scores nested within each of the four performance levels of the categorical variable. These are paraphrased as follows:

Level 1 – The student has not met the achievement standard and needs substantial improvement to demonstrate the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after high school.

Level 2 – The student has nearly met the achievement standard and may require further development to demonstrate the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after high school.

Level 3 – The student has met the achievement standard and demonstrates progress toward mastery of the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after completing high school coursework.

Level 4 – The student has exceeded the achievement standard and demonstrates the knowledge and skills in [content area] needed for likely success in entry-level credit-bearing college coursework after high school.

The percentages of these based on the frequencies against the total in each case are arranged in Table 10. For each content and grade level, these represent the summaries of the continuous field data, expressed in percentages, not distinguished by the teacher preparation variable.

Table 10

Percentages of Frequencies in Categorical Performance Levels across Grade and Content

Standard	Level 1 (Not Met)	Level 2 (Almost Met)	Level 3 (Met)	Level 4 (Exceeds)
English 5	16%	19%	28%	16%
English 11	28%	27%	30%	15%
Math 5	39%	30%	19%	12%
Math 11	54%	27%	11%	5%

These data in Table 10 show some considerable differences in the performance levels when comparing English and math teachers. At level 4 (exceeds standard), the data for English teachers indicated that their students exceeded the standard by almost a 2 to 1 margin when compared to those in math. A similar trend for English occurred at level 3 (met standard) by a 2 to 1 ratio. At the "not met" level, frequencies are 2 to 1 for math. Math 11 had the greatest percentage (58) for "not met" and the lowest percentage (5) for "exceeds." Conversely, English 5 had the lowest percentage for "not met." While significance was noted previously for traditionally trained teachers, it is interesting to observe that students in grades 5 and 11 in English posted more favorable frequencies than those in math, notwithstanding the significance for traditionally prepared teachers. However, it is also notable that all frequencies at the "exceeds" level averaged just 12% and 22% for "met." In all, about one-third of these students have achieved at an expected level.

Summary

Analysis of the data for both grade bands, 5^{th} and 11^{th}, indicated a statistically significant difference in rankings favoring teachers prepared in a traditional manner when compared to those prepared through an alternative pathway. When looking at the individual grade bands,

those teaching 5th grade English ranked significantly greater across all performance levels with a difference of 74 ranked scores compared to the traditionally prepared. Those teaching 5th grade mathematics showed a difference of 85 ranked scores. Teachers of 11th grade English showed the largest difference by 111 ranked scores for those traditionally prepared while those teaching 11th grade mathematics showed a difference of 43 ranked scores.

In effect, the overall results indicated that 5th and 11th grade WV traditionally prepared teachers ranked significantly greater on WV summative content exams than did their alternatively prepared peers. However, the data showed considerable differences in the performance levels when comparing English and math teachers. Data for traditionally prepared English teachers showed that their students exceeded the standard at a greater rate when compared to students taught by traditionally prepared math teachers.

Research Question #3:

"*What is the effect of teacher experience on West Virginia summative scaled scores in English language arts and mathematics among West Virginia students in grades 5 and 11?"*

English 5

Table 11

Descriptive Scaled Score Data for 5th Grade English, Teacher Experience Levels

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0-3 Years	4247	2490.79	87.433	1.342	2488.16	2493.42	2109	2743
4-9 Years	4245	2497.15	91.719	1.408	2494.39	2499.91	2177	2788
10> Years	9041	2504.76	89.333	.940	2502.91	2506.60	2175	2783
Total	17533	2499.53	89.647	.677	2498.21	2500.86	2109	2788

The data in Table 11 shows that the mean scaled scores modestly increased from those with the lesser experience to those with the greater experience.

Table 12

Analysis of Variance Data for 5th Grade English, Teacher Experience Levels

5th English	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	595149.613	2	297574.807	37.180	.000
Within Groups	140302619.397	17530	8003.572		
Total	140897769.011	17532			

The data in Table 12 indicated overall significance within the model (p .000 or < .0005), a multiple comparisons analysis showed significance (p .003) between all pairs of experience, i.e., between 0-3 and 4-9; between 0-3 and 10 > and between 4-9 and 10 >. In effect, as teacher experience increased, mean scaled scores increased respectively.

English 11

Table 13

Descriptive Data for 11th Grade English, Teacher Experience Levels

STUASMSCO

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0-3 YEARS	3953	2563.30	105.970	1.685	2559.99	2566.60	2102	2953
4-9 YEARS	5264	2571.54	109.048	1.503	2568.59	2574.48	2187	2880
10>YEARS	8958	2587.74	113.047	1.194	2585.40	2590.08	2102	2961
Total	18175	2577.73	110.859	.822	2576.12	2579.34	2102	2961

The same essentially held for 11th grade English teachers as it did for their 5th grade peers. The distributions of the scaled scores across the teacher experience categories increased respectively as the experience level increased. There is an overall difference of 25 scaled score points from bottom (0-3) to top (10>).

The variability is generally consistent and stable across the categories and the lower and upper bounds of the 95% confidence intervals include only a small range of values. Not unexpectedly, the greatest mean scaled scores occurred for those teachers with 10+ years of experience.

Table 14

Analysis of Variance Data for 11th Grade English Teacher Experience Levels

STUASMSCO

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1922866.110	2	961433.055	78.901	.000
Within Groups	221431541.770	18172	12185.315		
Total	223354407.879	18174			

The data in Table 14 indicated an overall significance within the model with a significant effect within the pairwise mean scores (p .000 or p< .0005). An effect size measure of .0625 indicated that approximately 6% of the total variability was accounted for by the predictor variable. Multiple comparisons, pairwise analysis resulted in significance between all pairs of experience, i.e., between 0-3 and 4-9; between 0-3 and 10> and between 4-9 and 10>. In effect, teacher experience progressively modulated scaled score effects.

Math 5

Table 15

Descriptive Data for 5th Grade Math, Teacher Experience Variable

STUASMSCO

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0-3 YEARS	4247	2490.79	87.433	1.342	2488.16	2493.42	2109	2743
4-9 YEARS	4245	2497.15	91.719	1.408	2494.39	2499.91	2177	2788
10> YEARS	9041	2504.76	89.333	.940	2502.91	2506.60	2175	2783
Total	17533	2499.53	89.647	.677	2498.21	2500.86	2109	2788

The results for grade 5 mathematics varied but still followed the trends previously established for years of experience. As the years of experience increased so did the mean scaled scores in each case. Additionally, the variability was stable and showed small interval values within the 95% confidence levels.

Table 16

Analysis of Variance for 5th Grade Math, Teacher Experience Variable

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	191452.131	2	95726.066	13.942	.000
Within Groups	115640288.667	16842	6866.185		
Total	115831740.799	16844			

The data in table 16 showed an overall test of significance for the model with p.000 (<.0005). Multiple comparisons indicated a significance for pairs 0-4 and 10>; between 4-9 and

10>; but not for 0-3 and 4-9. An effect size measure of .023 accounted for about 2% of the total variability. Figure 5 shows the points on the line for mean scaled scores per the experience nodes (0 = 0-4; 1 = 4-9 and 2 = 10> years).

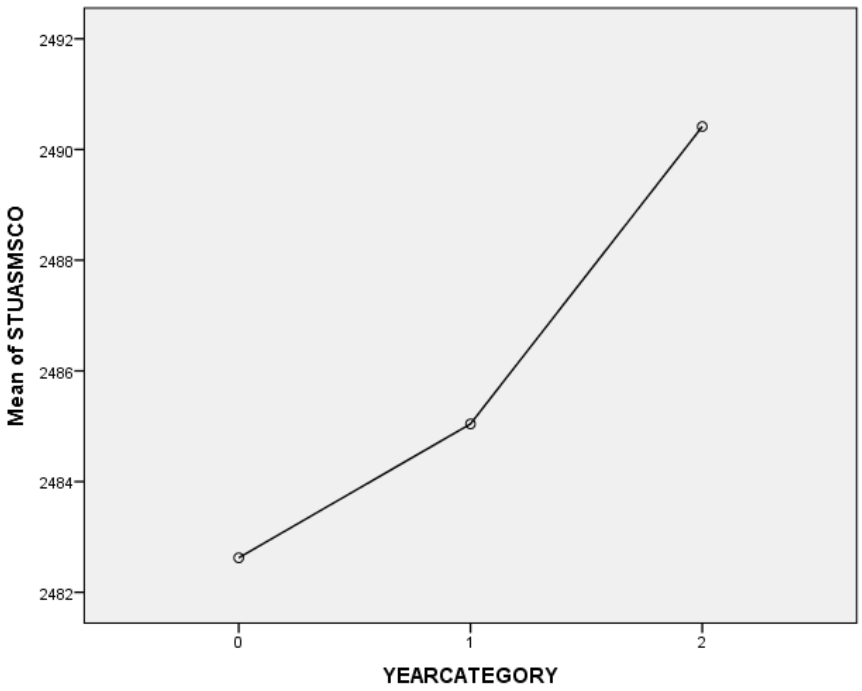


Figure 5. Line Plot of Experience Nodes for 0-3; -4-9 and 10> for Teacher Experience

Math 11

Table 17

Descriptive Data for 11th grade Math, Teacher Experience Levels

STUASMSCO

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0-3 Years	3057	2528.94	106.409	1.925	2525.17	2532.72	2118	2963
4-9 Years	3211	2524.02	117.465	2.073	2519.96	2528.08	2118	3085
10> Years	9949	2556.94	114.908	1.152	2554.68	2559.19	2118	3085
Total	16217	2545.14	114.841	.902	2543.37	2546.91	2118	3085

The previous trends described varied here—0-3 years had slightly greater scores than did 0-4 years. A difference of 25 scaled score points occurred from the bottom to the top of the experience levels. Again the greater score occurred for 10> years. Variability differed somewhat but remained within small values in the confidence intervals and differences in the standard deviations.

Math 11

Table 18

Analysis of Variance Data for 11th Grade Math, Teacher Experience Levels

STUASMSCO

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3618450.419	2	1809225.210	139.526	.000
Within Groups	210245961.380	16214	12966.940		
Total	213864411.799	16216			

Similarly, there is an overall significance (p.000 or p<.0005) with an effect size measure of .1699 or approximately 17% of the total variance accounted for. A multiple comparisons analysis resulted in significance between pairs 0-3 years and 10+ years (p.000) and between pairs 4-9 and 10> years (p .000). No significance occurred between 0-3 and 4-9 years (p >261). Figure 6 visually depicts these differences.

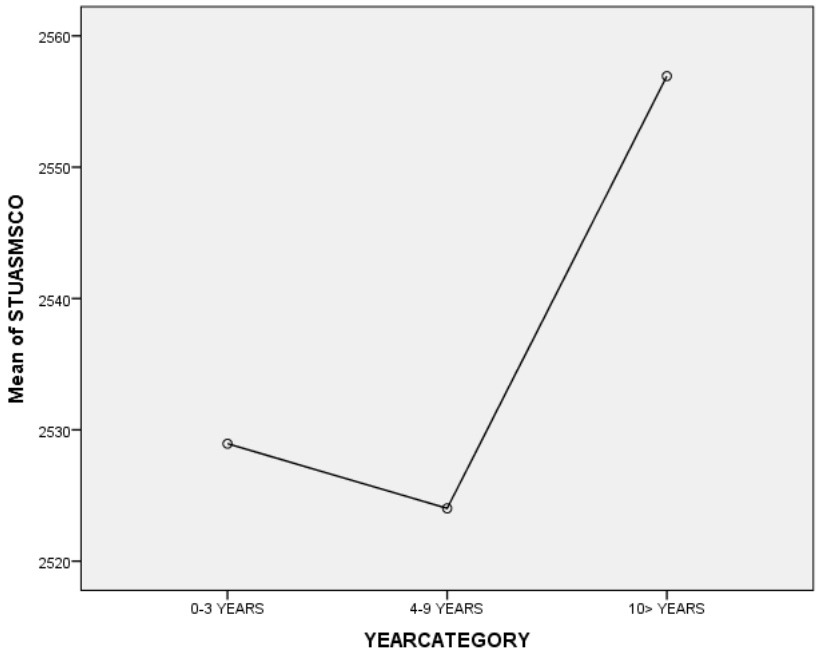


Figure 6. Teacher Experience Levels and Mean Scaled Scores for 11th grade Mathematics

Summary

Analysis of the data for grade bands 5th and 11th of those teaching English and mathematics indicated that there is a statistically significant difference across the teacher experience categories. The mean scaled scores increased respectively from those with the lesser experience. The variability is consistent across the categories. The greatest mean scaled scores occurred for those teachers with 10+ years of experience. When looking at those teaching 11th grade mathematics, the previous trends varied. Those with 0-3 years of experience had slightly

higher scores than those 0-4 years. Similarly, no significance in variance occurred between 0-3 and 4-9 years.

In effect, the overall results indicated that 5th and 11th grade WV students taught by more experienced teachers scored significantly greater on WV summative content exams than did their peers taught by less experienced teachers, except for 11th grade mathematics students.

Research Question #4:

"What is the effect of teacher experience on West Virginia Summative assessment categorical rankings (levels 1, 2, 3, and 4) for mathematics and English, grades 5 and 11?"

English 5

Table 19

Kruskal-Wallis Mean Ranks Data for Experience Levels of 5th Grade English Teachers

Ranks

YEARCATEGORY		N	Mean Rank
STUASMLEV	0-3 Years	4247	8347.70
	4-9 Years	4245	8657.15
	<10 Years	9041	9015.54
	Total	17533	

Data in Table 19 shows the mean ranks per the years of experience categories for 5th grade English teachers. In each case, the mean ranks increased respectively as the years of experience increased—with the greatest rank at 10+ years. The highest overall ranking occurred for 10+ years that corresponds to the highest score on the continuous variable.

Table 20

Kruskal-Wallis Test of Significance Data for 5th Grade English, Teacher Experience

Kruskal Wallace	STUASMLEV
Chi-Square	56.987
df	2
Asymp. Sig.	.000

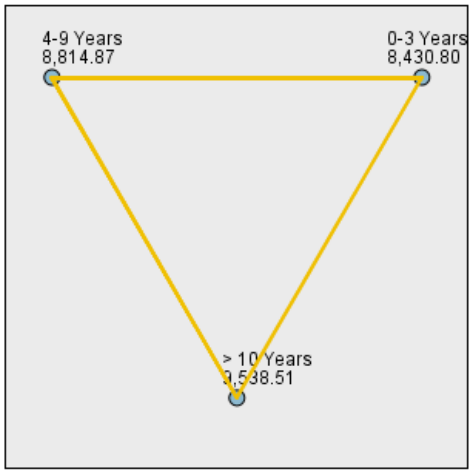
a. Kruskal Wallis Test
b. Grouping Variable: YEARCATEGORY

Data in Table 20 indicated a significant effect overall for the model with a p level of .000 or <.0005. In effect, there is a significant difference in the continuous variable across the three groupings, favoring teacher experience for 5th grade English teachers.

English 11

The output for 11th grade English was generated using the new module for Kruskal-Wallis, which combines visual and numerical output including pairwise comparisons of the three factors in the years of experience groupings. These data are shown in Figure 7.

Pairwise Comparisons of STAFFYEARS



Each node shows the sample average rank of STAFFYEARS.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
0-3 Years-4-9 Years	-384.068	106.394	-3.610	.000	.001
0-3 Years-> 10 Years	-1,107.715	96.528	-11.476	.000	.000
4-9 Years-> 10 Years	-723.648	87.793	-8.243	.000	.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 7. Visual and numerical data for staff years for 11th Grade English Teachers

The same effect occurred here for 11th grade English teachers. The test of significance for each of the nodes of experience showed a p level of .000 or < .0005. Multiple comparisons are visualized by the connecting gold lines in the staff year's portion of the figure. In each case, there is significance between the pairs. Likewise, the numerical data confirms the significance of the pairs in each case (p .000). In effect, the null hypothesis that the distribution of scores are the same along the categories of experience was rejected. All are significantly different, and significantly affected scaled scores as each level of teacher experience progressed.

Math 5

Table 21

Kruskal-Wallis Mean Ranks Data for Experience Levels of 5th Grade Mathematics Teachers

Ranks

YEARCATEGORY		N	Mean Rank
STUASMSCO	0-3 YEARS	3915	8160.10
	4-9 YEARS	4248	8328.00
	10 > YEARS	8682	8588.03
	Total	16845	

Table 21 shows the mean ranks per the years of experience categories. In each case, the mean ranks increased respectively as the years of experience increased—with the greatest rank at 10> years. The highest overall ranking occurred for 10+ years that corresponds to the highest score on the continuous variable. By observation of the mean ranks, it appears that there may be a difference between all three pairs of experience, considering a difference of 428 mean rank values from top to bottom.

Table 22

Kruskal-Wallis Test Statistics[a,b] of Years of Experience for 5th Grade Mathematics Teachers

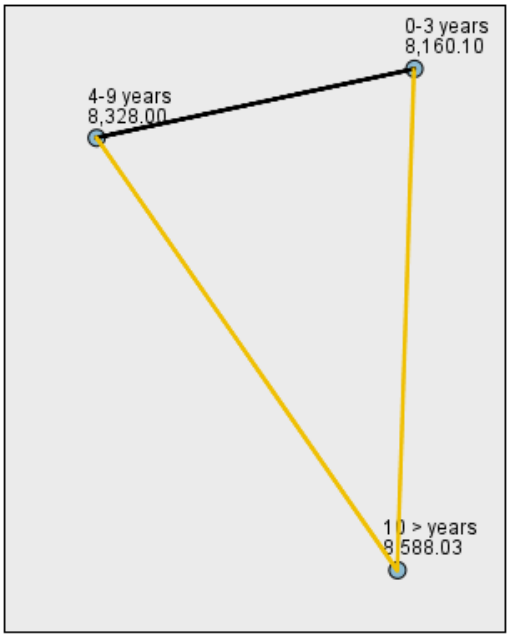
	STUASMSCO
Chi-Square	27.883
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test
b. Grouping Variable: YEARCATEGORY

Table 22 indicates a significant effect overall for the model with a p level of .000 or <.0005. In effect, there is a significant difference in the continuous variable across two or more of the three groupings.

Figure 8 identifies the significant pairs by the connecting gold lines. These lines show an effect or differences between 0-3 and 10+ years and 4-9 and 10+ years (p 000 (<.0005). However, no significance is apparent between 0-3 and 4-9 years. That outcome was confirmed by the test of significance and related p level shown in the numerical portion of the figure (p>103).

Pairwise Comparisons of YEARCATEGORY



Each node shows the sample average rank of YEARCATEGORY.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
0-3 years-4-9 years	-167.894	102.979	-1.630	.103	.309
0-3 years-10 > years	-427.930	89.483	-4.782	.000	.000
4-9 years-10 > years	-260.037	87.032	-2.988	.003	.008

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 8. Visual and Numerical Data for 5th Grade Math Teacher Experience

Math 11

Table 23

Mean Ranks Data for 11th Grade Math, Teacher Experience Variable

	STAFFYEARS	N	Mean Rank
STUASMLEV	0-3 Years	3057	7507.19
	4-9 Years	3211	7404.04
	> 10 Years	9949	8521.44
	Total	16217	

The data in Table 23 show the mean ranks per the years of experience categories for 11th grade math teachers. A difference of 117 rank values resulted from top to bottom rankings. The lowest rankings occurred for 4-9 years and the greatest for 10> years.

Table 24

Test Significance Data for 11th Grade Math, Teacher Experience Variable

Test Statistics[a,b]

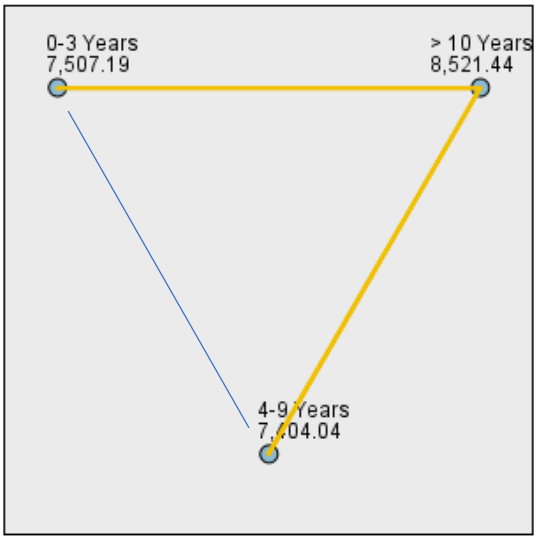
	STUASMLEV
Chi-Square	234.010
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test
b. Grouping Variable: STAFFYEARS

Table 24 indicates a significant effect overall for the model with a p level of .000 or <.0005. In effect, there is a significant difference in the continuous variable across two or more of the three groupings.

Figure 9 identifies the significant pairs for years of experience. Pairs between 0-3 and 10> and between 4-9 and 10> differed as shown by the connecting gold lines; however, pairs 0-3 and 4-9 years of experience were not connected significantly (blue line). That outcome was confirmed by the test of significance and related p level shown in the numerical portion of the

figure. Although the visual data shown in the ranks table showed a numerical difference of 103 ranked values, that effect did not occur statistically.



Each node shows the sample average rank of STAFFYEARS.

Sample1-Sample2	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
4-9 Years-0-3 Years	103.147	109.515	.942	.346	1.000
4-9 Years-> 10 Years	-1,117.395	87.962	-12.703	.000	.000
0-3 Years-> 10 Years	-1,014.248	89.621	-11.317	.000	.000

Each row tests the null hypothesis that the Sample 1 and Sample 2 distributions are the same.
Asymptotic significances (2-sided tests) are displayed. The significance level is .05.

Figure 9. Visual and Numerical Data for Experience Nodes of 11th Grade Math Teachers

Summary

Analysis of the data for grade bands 5 and 11 of those teaching English indicated a statistically significant difference. The mean ranks increased respectively as the years of experience increased. Once again, the greatest rank occurred for those with 10+ years of experience. Data also showed that there is a significant difference in the continuous variable

across the three groupings (0-3, 4-9, <10) favoring teacher experience for 5th grade English teachers. When comparing teachers of mathematics for both 5th and 11th grades, no significance occurred between 0-3 and 4-9 years.

In effect, the overall results indicated that 5th and 11th grade English language arts WV students taught by more experienced teachers ranked significantly greater on WV summative content exams than did their peers taught by less experienced teachers, except for 5th and 11th grade mathematics students. A great difference of 1,117 ranked values resulted between 0-4 and 10> years and 1,114 ranked values between 0-3 and 10> years.

Research Question #5:

"*What Is The Effect Of The Type Of Teacher Licensure On West Virginia Summative Scaled Scores In English and Mathematics among West Virginia Students In Grades 5 And 11?*"

English 5

Table 25

Descriptive Data for 5th Grade English Teachers, Certified Variable

STUASMSCO

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Not Certified	287	2387.69	75.656	4.466	2378.90	2396.48	2223	2645
Certified	364	2504.45	81.947	4.295	2496.00	2512.90	2269	2711
Total	651	2452.98	98.155	3.847	2445.42	2460.53	2223	2711

Table 25 shows the means scores favoring those certified by a difference of approximately 117 scaled score points. Variability in the model was considered homogeneous, with a minor difference of less than 2%. Sample sizes for the groupings were essentially equivalent.

Table 26

Analysis of Variance for 5th Grade English Teachers, Certified Variable

STUASMSCO

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2187738.096	1	2187738.096	348.455	.000
Within Groups	4074673.511	649	6278.388		
Total	6262411.607	650			

Table 26 indicates an overall (omnibus) test of significance for the model, with a p level of .000 or <.0005, again favoring those certified. An effect size measure of .349 indicated that approximately 35% of the total variance in the model was accounted for by the predictor variable.

Math 5

Table 27

Descriptive Data for 5th Grade Math Teachers, Certified Variable

STUASMSCO

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0	242	2375.83	72.020	4.630	2366.72	2384.95	2223	2619
1	348	2486.73	80.855	4.334	2478.20	2495.25	2266	2776
Total	590	2441.24	94.625	3.896	2433.59	2448.89	2223	2776

Table 27 shows the means scores for 5th grade math certified and noncertified teachers—favoring certified teachers by a difference of approximately 111 scaled score points. Variability in the model was considered homogeneous, with a minor difference of less than 3% for the standard error.

Table 28

Analysis of Variance for 5th Grade Math Teachers, Certified Variable

STUASMSCO **Anova**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1755275.886	1	1755275.886	293.332	.000
Within Groups	3518546.455	588	5983.923		
Total	5273822.341	589			

The data in Table 28 indicate an overall (omnibus) test of significance for the model, with p level of .000 or <.0005, again favoring those certified. The effect size is .2146, which indicates that approximately 22% of the total variance in the model was accounted for by the predictor variable.

Math 11

Table 29

Descriptive Data for 11th Grade Math Teachers, Certified Variable

STUASMSCO

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Not Certified	797	2413.05	96.606	3.422	2406.33	2419.77	2118	2921
Certified	642	2548.01	105.242	4.154	2539.85	2556.16	2245	2894
Total	1439	2473.26	120.860	3.186	2467.01	2479.51	2118	2921

Table 29 shows the means scores for those certified and noncertified for 11th grade math teachers. Those certified had an average scaled score of 2,548 compared to 2,413 for those noncertified, favoring certified by a difference of approximately 135 scaled score points. Variability in the model was considered homogeneous, with minor differences in the values for standard deviations considering the mean score values.

Table 30

Analysis of Variance for 11th Grade Math Teachers, Certified Variable

STUASMSCO **Anova**

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6476424.744	1	6476424.744	640.575	.000
Within Groups	14528544.053	1437	10110.330		
Total	21004968.796	1438			

Data in Table 30 indicate an overall (omnibus) test of significance for the model, with p level of .000 or <.0005, again favoring those 11th grade math teachers who were certified. The effect size is .3083, which means that approximately 31% of the total variance in the model is accounted for by the predictor variable. In effect, certified 11th grade math teachers had students with significantly greater mathematics scaled scores than did those students taught by noncertified math teachers.

English 11

Table 31

Descriptive Data for 11th Grade English Teachers, Certified Variable

Group Statistics

CERTIFIED		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	Not Certified	3005	2568.49	122.064	2.227
	Certified	3218	2574.82	110.309	1.945

Table 31 shows the means scores for those certified and noncertified for 11th grade English teachers. Those certified had an average scaled score of 3,218 compared to 3,005 for those noncertified, favoring those certified by a difference of approximately 213 scaled score points. Variability in the model was considered homogeneous, with minor differences in the values for standard deviations considering the mean score values.

Table 32

Independent Samples T-Test of Significance for 11th Grade English Teachers

Independent Samples Test

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	46.977	.000	2.149	6221	.032	-6.331	2.946	12.106	-.556
Equal variances not assumed			-2.12	6048.218	.032	-6.331	2.956	12.126	-.536

The data in Table 32 show an overall (omnibus) test of significance for the model, with a p level of .000 or <.0005, again favoring those 11th grade English teachers who were certified. An effect size measure of .3083 meant that approximately 31% of the total variance in the model was accounted for by the predictor variable. In effect, certified 11th grade English teachers taught students with significantly greater English scaled scores than did those students taught by noncertified English 11th grade teachers.

Summary

Analysis of the data for grade bands 5 and for English and mathematics teachers indicated a statistically significant difference favoring certified teachers when compared to those noncertified. Eleventh grade English teachers showed the biggest difference with a difference of 213 scaled score points for the certified teacher. The next most significant difference occurred in 11th grade mathematics. Certified teachers in this area showed a difference of approximately 135 scaled score points when compared to their noncertified peers. Those certified and teaching 5th grade English showed a difference of approximately 117 scaled score points. Certified 5th grade mathematics teachers showed a difference of approximately 111 scaled score when compared to noncertified peers.

In effect, the overall results indicated that WV students in grades 5 and 11 who were taught by certified teachers scored significantly greater on WV summative content exams than did their peers taught by noncertified teachers, notwithstanding the content area.

Research Question #6:

"*What Is the Effect of Type of Licensure (Certified/NonCertified) on Categorical Rankings in Math and English, Grades 5 and 11?"*

English 5

Table 33

Mean Ranks English 5

Ranks

	CERTIFIED	N	Mean Rank	Sum of Ranks
STUASMLEV	Not Certified	287	213.56	61293.00
	Certified	364	414.65	150933.00
	Total	651		

These data were analyzed by the Mann-Whitney U Median Ranks (MWU) model, which compared differences in the distributions of mean ranks for certified and noncertified 5th grade English teachers. The data in Table 33 show that certified teachers ranked significantly greater across the four categorical levels for English 5, with a difference of 201 ranked scores compared to those noncertified.

Table 34

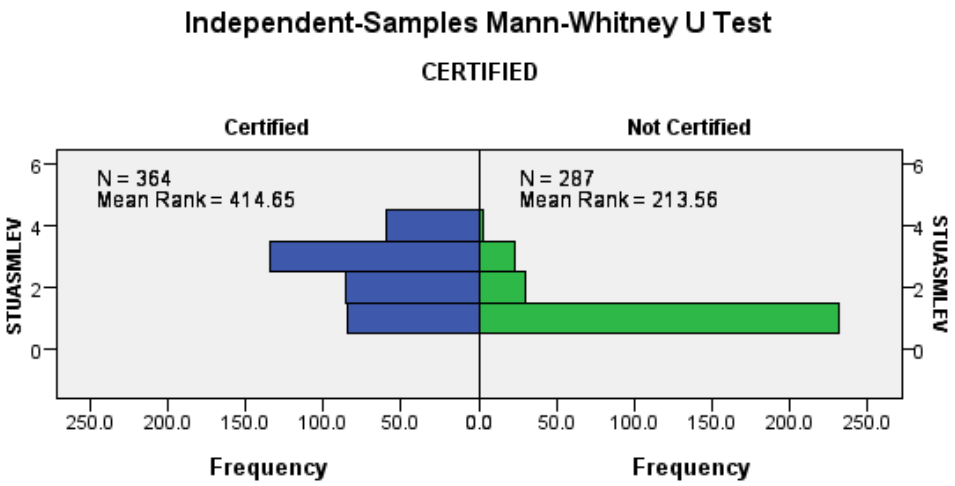
Group Statistics English 5

Test Statistics[a]

	STUASMLEV
Mann-Whitney U	19965.000
Wilcoxon W	61293.000
Z	-14.554
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: CERTIFIED

Data in Table 34 confirms statistical significance with a *z* approximation test score of 14.554 and a test value of p .000 or <.0005. In effect, there is a statistically significant difference in ranked scores favoring the direction of certified teachers. An effect size measure of approximately .1025 indicated that 10% of the total variability was accounted for by the predictor variable (Certified). Figure 10 visually and numerically depicts these outcomes and confirms the test of significance. Additionally, it is notable that nearly 250 5th graders taught by noncertified teachers placed within Level 1 of the ranks data (standard not met*)*. This compared to approximately 75 who were taught by certified teachers.



Total N	651
Mann-Whitney U	84,503.000
Wilcoxon W	150,933.000
Test Statistic	84,503.000
Standard Error	2,217.249
Standardized Test Statistic	14.554
Asymptotic Sig. (2-sided test)	.000

Figure 10. Median Ranks Data for Effects on Type of Licensure for Certified and Noncertified 5th Grade English Teachers

English 11

Table 35

Group Statistics Data for Certified and Noncertified 11th Grade English Teachers Group Statistics

CERTIFIED		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	Not Certified	3005	2568.49	122.064	2.227
	Certified	3218	2574.82	110.309	1.945

The data in Table 35 compares the mean scaled scores for certified and noncertified 11th grade English teachers. A difference resulted of approximately 7 rank points when comparing the mean scaled scores for certified and noncertified 11th grade English teachers. Variability was essentially equivalent given the minor differences in standard deviation and standard error of mean values. Sample sizes were essentially equivalent.

Table 36

Mann-Whitney U Test of Independence for Certified and Noncertified 11th Grade English Teachers Across the Categorical Levels

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	46.977	.000	2.149	6221	.320	-6.331	2.946	12.106	-.556
Equal variances not assumed			2.142	6048.218	.320	-6.331	2.956	12.126	-.536

The Mann-Whitney U model was calculated to test whether two independent samples (certified and noncertified) are from the same distribution or differed significantly across the categorical levels. The data in Table 36 indicated no significance for scaled scores among 11th grade certified and noncertified English teachers. Based on the test of significance and a p level of .320 (>.05) there is no statistical difference related to one's certification status for 11th grade English teachers across the categorical levels.

Math 5

Table 37

Mean Ranks Data for Certified and Noncertified 5th Grade Math Teachers

Group Statistics

CERTIFIED STUASMLEV	N	Mean Rank	Sum of Ranks
Not Certified	242	203.45	49234.00
Certified	348	359.51	125111.00
Total	590		

The data in Table 37 show the mean scaled scores for certified and noncertified 5th grade math teachers and related variability. A difference resulted of approximately 55 mean ranked points when comparing the mean scaled scores for certified and noncertified 5th grade Math teachers across the categorical levels, which favored certified teachers. Variability was essentially equivalent given the minor differences in standard deviation and standard error of mean values. Sample sizes were essentially equivalent.

Table 38

Mann-Whitney U Test of Independence for Certified and Noncertified 5th Grade Math Teachers

Test Statistics[a]

	STUASMLEV
Mann-Whitney U	19831.000
Wilcoxon W	49234.000
Z	-12.325
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: CERTIFIED

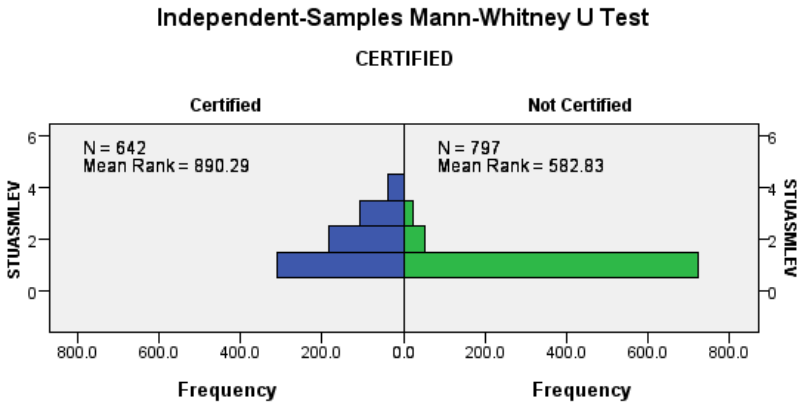
The Mann-Whitney U "Legacy" model was used to test whether two independent samples (certified and noncertified) are from the same distribution or differed significantly

across the categorical levels. The data in Table 38 indicated statistical significance for scaled scores among 5th grade certified and noncertified Math teachers (z, 12.325; p .000, or < .0005.). An effect size measure of 15.4 indicated that approximately 15% of the total variability was accounted for by the predictor variable (certified).

Math 11

The Mann-Whitney U New Module (MWU) was used to calculate the data for 11th grade Math teachers. The new MWU model on SPSS (Version 24) produces a combination of numerical and graphic outcomes which displays descriptive statistical and visual data as well as tests of statistical significance. These data are seen in Figure 11.

Data showed a mean rank of 583 for those noncertified compared to a mean rank of 890 for those certified, resulting in a difference of 313 ranked points favoring 11th grade certified mathematics teachers. Differences in frequencies of students within each of the categorical levels are also given. Interestingly, there are about 260 students at level 1 associated with certified teachers (standard not met) compared to 780 at the same level associated with noncertified teachers. In effect, approximately two-thirds (66%) of the 11th grade students taught mathematics by noncertified teachers do not meet mathematics standards. Tests of statistical significance for the data shown in Figure 11 confirm the considerable difference in rankings (Z, 17.648; p .000 or < .0005).



Total N	1,439
Mann-Whitney U	365,161.000
Wilcoxon W	571,564.000
Test Statistic	365,161.000
Standard Error	6,194.579
Standardized Test Statistic	17.648
Asymptotic Sig. (2-sided test)	.000

Figure 11. Mann-Whitney U Numerical and Graphic Data for Certified and Noncertified 5th Grade Math Teachers

Summary

Analysis of the data for grade bands 5 and 11 for mathematics and for 5th grade English indicated a statistically significant difference in rankings favoring certified teachers. Certified 5th grade English teachers ranked significantly greater across the four categorical levels with a difference of 201 ranked scores compared to those noncertified. Those certified and teaching 5th grade mathematics showed a difference of approximately 55 mean ranked points when compared to those noncertified in the same grade band. The same held true for 11th grade mathematics with a difference of 313 ranked points favoring certified teachers. Of note, approximately two-thirds (66%) of the 11th grade students taught mathematics by noncertified teachers do not meet the

mathematic achievement standard for that grade level. In contrast, data for 11th grade English showed no significance for scaled scores among certified and noncertified teachers.

In effect, the overall results indicated that for 5th and 11th grade mathematics and for 5th grade English, WV students taught by certified teachers ranked significantly greater on WV summative content exams than did their peers taught by noncertified teachers. An exception was for 11th grade English students where teacher certification showed no statistically significant difference.

Research Question #7:

"Effect on Math and English Scaled Scores for 5th Grade and 11th Grade Math and English Teachers with Highly Qualified Status or Non-highly Qualified Status."

English 5

Table 39

Descriptive Data for Highly and Non-highly Qualified 5th Grade English Teachers

Group Statistics

HQENGMATH		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	0 not highly qualified	108	2427.59	94.933	9.135
	1 highly qualified	120	2461.68	95.790	8.744

The data in Table 39 include mean scaled scores and related variability measures for 5th graders taught by English teachers with highly and non-highly qualified status. Numerical data shows a difference of approximately 34 scaled score points for those 5th grade students taught by highly qualified English teachers. Variability appears to be homogeneous considering the minor

differences in standard deviations and standard error of the mean values. Sample sizes are essentially equal.

Table 40

Independent Samples T-Test for Highly Qualified and Non-highly Qualified 5th Grade English Teachers

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	.362	.548	2.960	649	.003	-30.431	10.281	50.618	10.244
Equal variances not assumed			3.026	155.876	.003	-30.431	10.058	50.299	10.564

The scaled scores were analyzed by an Independent Samples T-Test to test the significance of the teacher status variable. The data in Table 40 show a statistical significance of p .003, (< .005) which favored highly qualified English teachers. Overall, highly qualified 5th grade English teachers had a greater effect on English scaled scores of their students than did their non-highly qualified peers. An effect size of .054 indicated that approximately 5% of the total variance in the study was accounted for by the predictor variable (quality) which is a small

effect size according to Cohen's guidelines. Guidelines for interpreting effect size values provided by Cohen (1992) are: (.20 small effect; .50 moderate effect and .80> large effect).

English 11

Table 41

Independent Samples T-Test for Highly Qualified and Non-highly Qualified 11th Grade English Teachers

Group Statistics

HQENGMATH		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	0	737	2560.45	125.968	4.640
	1	814	2574.35	111.001	3.891

Table 41 includes mean scaled scores and related variabilities for 11th grade students taught by highly qualified and non-highly qualified 11th grade English teachers. A difference of 14 mean scaled score points favored students who were taught by highly qualified 11th grade English teachers. Variability appears to be homogeneous with a minimal difference in standard deviation and standard error of the mean values. Sample sizes are essentially equivalent.

Table 42

Independent Samples T-Test for Highly Qualified and Non-highly Qualified 11th Grade English Teachers

Independent Samples Test

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	18.002	.000	2.310	1549	.021	-13.904	6.018	25.707	-2.00
Equal variances not assumed			2.296	1474.764	.022	-13.904	6.055	25.782	2.026

Mean scaled scores were analyzed by an Independent Samples T-Test to determine the significance of the teacher quality variable. A comparative analysis of the two sets of scores in the status grouping show a statistical significance of p .022, (< .05) which favored the mean score for highly qualified 11th grade English teachers. Overall, these English teachers had the greater effect on English scaled scores of their students than did their non-highly qualified peers. An effect size of 12.2 indicated that approximately 12% of the total variance in the study was accounted for by the predictor variable (quality). The effect size is considered to be a small effect (Cohen, 1992).

Math 5

Table 43

Group Statistics Data for Highly Qualified and Non-highly Qualified 5th Grade Math Teachers

Group Statistics

HQENGMATH		N	Mean	Std. Deviation	Std. Error Mean
STUASMSCO	0	95	2420.94	94.728	9.719
	1	115	2455.12	91.809	8.561

The mean score data in Table 43 show a difference of 34 mean scaled score points favoring highly qualified 5th grade math teachers. Variability appears to be homogeneous considering the minor differences in values for standard deviation and standard error of the mean. Sample sizes are essentially equivalent.

Table 44

Independent Samples T-Test for Highly Qualified and Non-highly Qualified 5th Grade English Teachers

Independent Samples Test

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	.005	.944	2.647	208	.009	-34.185	12.913	59.642	-8.727
Equal variances not assumed			2.639	198.118	.009	-34.185	12.952	59.726	-8.644

Mean scaled scores were analyzed by an Independent Samples T-Test to determine the significance of the teacher quality variable. A comparative analysis of the two sets of scores in the status grouping shows a statistical significance of p .009 (< .005) which favored the mean score for highly qualified 5th grade math teachers. Overall, highly qualified 5th grade math teachers had the greater effect on the mean scaled scores of their students than did their non-highly qualified peers. An effect size of .067 indicated that approximately 7% of the total variance in the study was accounted for by the predictor variable (quality). The effect size is considered to be a small effect (Cohen, 1992).

Math 11

Table 45

Group Statistics Data for Highly Qualified and Non-highly Qualified 11th Grade Math Teachers

Group Statistics

HQENGMATH STUASMSCO	N	Mean	Std. Deviation	Std. Error Mean
0	460	2454.66	112.604	5.250
1	487	2482.07	122.877	5.568

The data in Table 45 show the mean scaled scores and related variability for 11th grade math students taught either by highly qualified or non-highly qualified teachers. A difference of approximately 27 mean scaled score was found for those students taught by highly qualified Math 11 teachers compared to their peers taught by non-highly qualified teachers. Variability appears to be homogeneous considering the minor differences shown for standard deviations and the standard error for the mean. Sample sizes are essentially equivalent.

Table 46

Independent Samples T-Test of Scaled Scores for Highly Qualified and Non-highly Qualified 11th Grade Math Teachers

	Levene's Test for Equality of Variances		T-Test for Equality of Means						
								95% Confidence Interval of the Difference	
STUASMSCO	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	5.893	.015	3.574	945	.000	-27.417	7.672	42.474	12.361
Equal variances not assumed			3.583	944.141	.000	-27.417	7.653	42.436	12.399

Mean scaled score data in Table 46 were analyzed by an Independent Samples T-Test to test the significance of the teacher quality variable for the two groupings. A comparative analysis of the two sets of scores in the status grouping showed a statistical significance of p .000 (< .0005) which favored the mean scores of students taught by highly qualified 11th grade mathematics teachers. Overall, highly qualified 11th grade mathematics teachers had the greater effect on scaled scores of their students than did their non-highly qualified peers. An effect size of .131 indicated that approximately 13% of the total variance in the study was accounted for by the predictor variable (quality). The effect size is considered to be a small effect (Cohen, 1992).

Summary

The results for 5th and 11th grade bands for mathematics and English indicated a statistically significant difference in scaled scores favoring highly qualified teachers. Students taught by highly qualified English teachers outscored their peers by 34 mean scaled score points than students who were taught by non-highly qualified teachers. Eleventh grade English students taught by highly qualified teachers showed a difference of 14 mean scaled score points when compared to those non-highly qualified. The same trend held true for 5th and 11th grade mathematics with a difference of 34 mean scaled score points favoring highly qualified teachers for 5th grade Mathematics and 27 mean scaled score points for 11th grade mathematics.

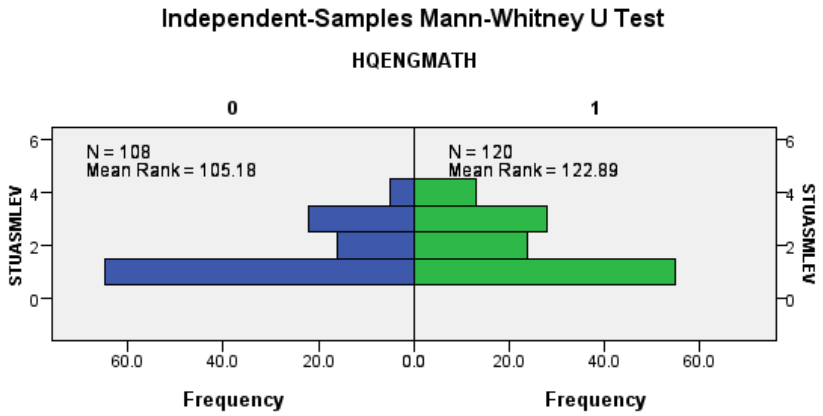
In effect, the overall results indicated that 5th and 11th grade English and mathematics students taught by highly qualified teachers scored significantly greater on WV summative content exams than did their peers taught by non-highly qualified teachers.

Research Question #8:

"What is the Effect of Highly and Non-highly Qualified Teacher Status on Categorical Rankings in English and Mathematics for Grade 5 and 11?"

English 5

The Mann-Whitney U module (MWU) was used to calculate the data for English 5. The new MWU model on SPSS (Version 24) produces a combination of numerical and graphic outcomes which displays descriptive statistical and visual data as well as inferential tests of statistical significance. These data are seen in Figure 12.



Total N	228
Mann-Whitney U	7,487.000
Wilcoxon W	14,747.000
Test Statistic	7,487.000
Standard Error	455.193
Standardized Test Statistic	2.212
Asymptotic Sig. (2-sided test)	.027

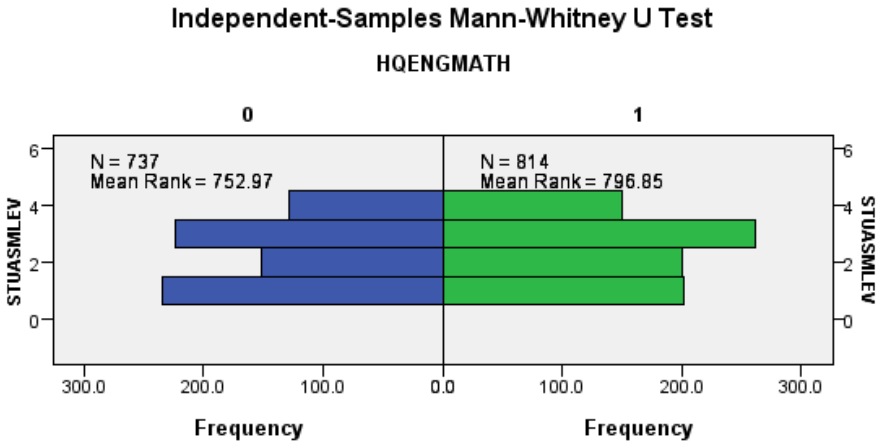
Figure 12. Mann-Whitney U Median Ranks Descriptive and Inferential Data for Effects on Categorical Rankings Among Qualified and Nonqualified 5th Grade English Teachers

Data analyzed showed a mean rank of 123 for those 5th grade English teachers who were highly qualified compared to a rank of 105 for those not highly qualified. A result of 18 ranked scores favored highly qualified 5th grade English teachers. Differences in frequencies of students within each of the categorical levels were also given. Notably, there were about 58 students at level 1 (standard not met) taught by highly qualified teachers compared to 63 at the same level associated with non-highly qualified teachers. In this instance, the effects of categorical rankings for English 5 were not distinguished by the teacher quality variable. Tests of statistical

significance for the data shown in Figure 12 confirm the overall significance in rankings *(z,* 2.212; p .027 or < .05).

English 11

Data for 11th grade English students and teachers likewise were analyzed by the Mann-Whitney U new module. The analysis tested the hypothesis that the distributions (median ranks) for non-highly qualified and highly qualified 11th grade English teachers were not equivalent. These results are shown in Figure 13.



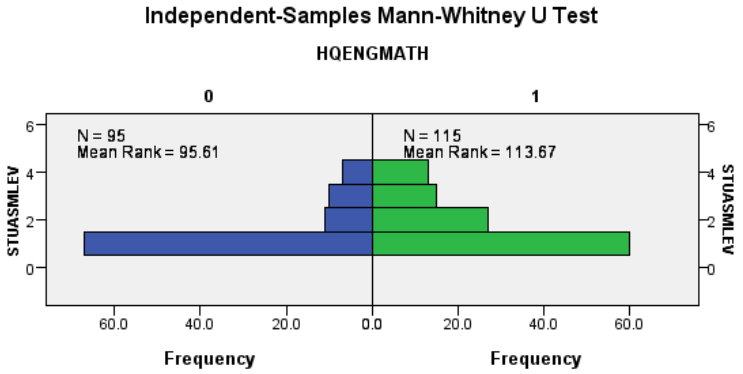
Total N	1,551
Mann-Whitney U	316,929.000
Wilcoxon W	648,634.000
Test Statistic	316,929.000
Standard Error	8,493.508
Standardized Test Statistic	1.998
Asymptotic Sig. (2-sided test)	.046

Figure 13. Mann-Whitney U Median Ranks Descriptive and Inferential Data for Effects on Categorical Rankings Among Qualified and Nonqualified 11th Grade English Teachers

The visual data in the population pyramid for the grouping distributions appeared to be very similar in regard to the frequencies in the ranks. The numerical data showed a difference of approximately 42 mean rank scores favoring highly qualified 11th grade English teachers, which was statistically significant with a p level of .046 (<.05). In effect, the differences in mean ranks were distinguished by the teacher quality variable.

Math 5

Data for 5th grade math teachers and students likewise were analyzed by the Mann-Whitney U new module. The analysis tested the hypothesis that the distributions (median ranks) for non-highly qualified and highly qualified 5th grade English teachers were not equivalent. These results are shown in Figure 14.



Total N	210
Mann-Whitney U	6,402.000
Wilcoxon W	13,072.000
Test Statistic	6,402.000
Standard Error	384.687
Standardized Test Statistic	2.442
Asymptotic Sig. (2-sided test)	.015

Figure 14. Mann-Whitney U Median Ranks Descriptive and Inferential Data for Effects on Categorical Rankings Among Qualified and Nonqualified 5th Grade Math Teachers

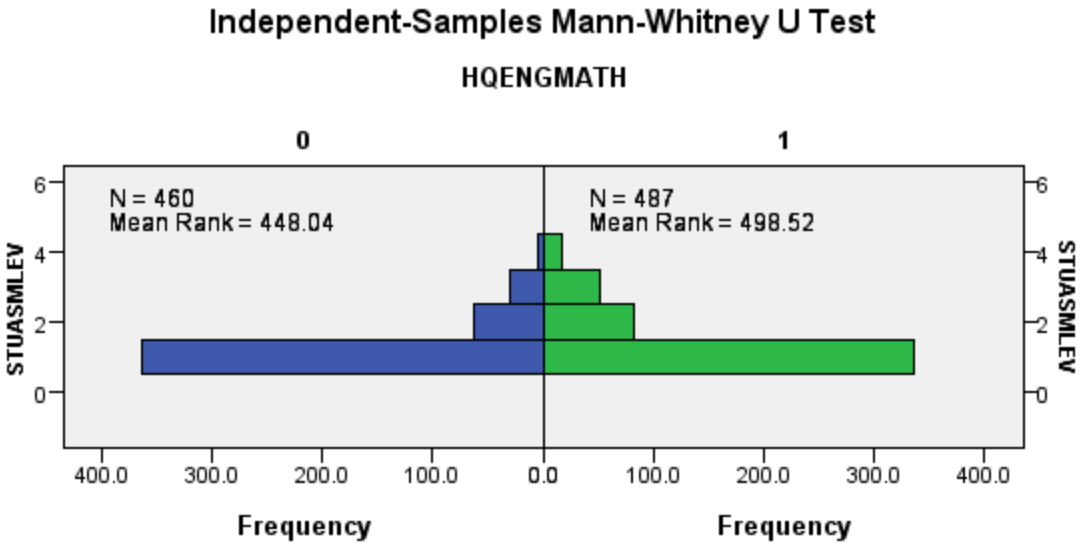
These results show a difference of 18 mean rank scores in the direction for highly qualified 5th grade math teachers. That difference was large enough to claim statistical significance with a p level of .015 (< .05) and a z approximation score of 2.442. However, it is observable in the ranks table of Figure 14 that the frequencies for level 1 (standard not met) were considerably larger for both groupings in comparison to the other three levels.

Notwithstanding the significance found for highly qualified 5th grade math teachers, it is observable that the frequencies for level 1 (standard not met) were essentially equivalent, i.e., the teacher quality variable did not particularly impact the achievement of students for meeting 5th grade math standards.

Math 11

These data were also analyzed by the Mann-Whitney U, independent samples new module. The analysis tested the hypothesis that the distributions (median ranks) for non-highly qualified and highly qualified 5th grade math teachers were not equivalent. These results are shown in Figure 13. These data show a difference of 51 mean rank scores in the direction of highly qualified 11th grade math teachers. A standardized test significance (*z* score) of 3.683 and a related p level of .000 (<.0005) confirmed that difference.

Figure 15 also shows the overall distributions for the scaled ranking performance levels for 11th graders in math. It is observable in the population pyramid that the frequencies for level 1 (standard not met) were essentially equivalent (60 cases for highly qualified and about 65 for non-highly qualified). Notwithstanding, the overall significance in the model for highly qualified teachers, the teacher quality variable did not make a significant impact on reducing the ranks for "standard not met."



Total N	947
Mann-Whitney U	123,949.500
Wilcoxon W	242,777.500
Test Statistic	123,949.500
Standard Error	3,241.401
Standardized Test Statistic	3.683
Asymptotic Sig. (2-sided test)	.000

Figure 15. Mann-Whitney U Median Ranks Descriptive and Inferential Data for Effects on Categorical Rankings Among Qualified and Nonqualified 11th Grade Math Teachers

Summary

Analysis of the data for 5th grade English and mathematics, and 11th grade mathematics indicated a statistically significant difference in scaled rankings favoring highly qualified teachers when compared to those non-highly qualified. Fifth grade English students taught by

highly qualified teachers had higher rankings scores by a difference of 18 mean rank scores compared to those of non-highly qualified. Fifth grade mathematics students taught by highly qualified teachers showed a difference of 18 mean rank scores when compared to those non-highly qualified. The same held true for 11th grade mathematics with a difference of 51 mean rank scores favoring highly qualified. Of note, for 5th and 11th grade mathematics, there was no statistically significant difference for students who ranked in level 1 (standard not met). Furthermore, for 11th grade English teachers, that difference was not statistically significant. The differences in mean ranks were not distinguished by the teacher quality variable.

In effect, the overall results indicated that 5th grade English and 5th and 11th grade mathematics students taught by highly qualified teachers ranked significantly greater on WV summative content exams than did their peers taught by non-highly qualified teachers with the exception of the those scoring in level 1.

CHAPTER 5

CONCLUSIONS, DISCUSSION AND RECOMMENDATIONS

Overview

The purpose of the study was to determine the effect of traditional and alternative teacher preparation, years of service, type of licensure held and teacher quality on West Virginia students' English language arts and mathematics summative assessment scaled scores and performance levels in grades 5 and 11.

This study collected licensure, employment, and assessment data to examine the level of impact on grade level standardized tests in mathematics and English language arts moderated by the kind of teacher preparation, years of teacher service, type of licensure held and highly qualified status. Its initial design is a causal-comparative, post hoc, non-equivalent model. The major outcome variables were English language arts and mathematics scaled scores and related categorical rankings on the summative, year-end assessment in West Virginia.

The study included all mathematics and English teachers who were currently employed in a WV public school in all 55 counties during academic year 2015-2016 in grades 5 and 11. These subjects were obtained from an encrypted statewide database at the West Virginia Department of Education (WVDE).

There were two sub-populations in the design: all 3,589 mathematics and English classroom teachers in grade levels 5 and 11 in West Virginia's 55 counties during the 2015-16 school year, and their respective numbers of English/language arts and mathematics students in grades 5 and 11 for a total of 34,528. However, 155 teachers were omitted because their preparation pathway could not be clearly identified as traditional or alternative. In all, there were 3,434 included in the population. Within the sample, there is an overlap in the variables and

factors of the study. For example, teacher experience encompasses all participants in the study no matter the variable.

Research Questions

The study addressed eight research questions designed to determine the effects on WV summative test scores and categorical rankings in English and mathematics among 5th and 11th graders moderated by their teachers': years of classroom teaching experience; type of teacher certification; type of teacher preparation program; and highly qualified status.

Categorical variables were likewise assessed which included four levels of student performance in regard to WV standards as distinguished by these same variables.

Findings

1. Highly qualified mathematics teachers had a very positive impact on their students' math achievement in grades 5 and 11.
2. As the years of classroom teaching experience increased, it was likely that achievement for 5th and 11th graders in mathematics increased as well as for English language arts.
3. Having teachers with professional teacher certification made a positive impact on 5th and 11th grade student summative test scores in math and English.
4. It appears that the achievement of 5th and 11th graders was benefitted when they were taught by teachers trained in traditional teacher preparation programs. This held true for English and mathematics.
5. The research variables in the study were positively related to student performance across the 4 performance levels and the related standards (exceeds, has met, nearly met, and not met).

6. Although significance occurred overall for the categorical rankings in regard to the variables, no differences occurred for 11th grade English teachers.
7. Differences occurred in the frequencies of students whose mean scaled scores placed them in the 4 categorical performance levels. A ratio of 5 to 1 students placed at the "standard not met" who were taught by noncertified math teachers.
8. Conversely, a ratio of about 1 to 1 occurred at level 1 (standard not met) for frequencies related to the teacher quality variable.
9. Overall, frequencies of students across the categorical rankings were exceptionally greater for the "standard not met." This held true for the teacher preparation, teacher certification and teacher experience variables.
10. Notwithstanding the statistical significance described for the study variables, these effects did not make significant inroads into the exceptionally large number of frequencies for "standard not met."

Conclusions And Discussion

Allowing persons to enter the teaching profession through various alternative certification pathways has mixed results nationally about the effectiveness of these programs on student learning and achievement. One of the reasons is because of the great variability in their requirements and contexts and the evolution of a great many of these programs since the 1980s. The effects of these programs also vary with some research supporting and some contesting their effectiveness. Redding and Smith (2016) indicate no differences on student effects between alternative certification pathways and traditional preparation programs; Allen (2003) states that alternatively prepared teachers are as effective as traditionally prepared teachers; Fowler (2003) found no differences in quality between alternatively prepared and traditionally prepared teachers; Sass (2013) noted that alternatively prepared teachers have greater involvement with minorities. Decker, Mayer and Glazerman (2004) concurred that alternatively prepared teachers make significant impact on mathematics achievement.

A report by the U.S. Secretary of education on alternative certification programs claimed that traditional teacher preparation programs are weak and that alternative certification programs attract academically stronger students. Moreover, these teachers significantly improve student achievement (Educational Research Newsletter, 2003). Critics of the report pointed out that, of the 44 studies reviewed, only one was evaluated by a blind peer review board. In effect, the data were not considered to be scientifically investigated. In the current study certain controls were implemented including random sampling and assignment, large, statewide sample sizes, various inferential techniques and effect size measures.

Additional evidence regarding impact of alternative certification programs and traditional preparation programs is documented by Ingersoll, Merrill, and May (2014) who conclude that

traditional programs directly and significantly impact student achievement; by Darling-Hammond (2007) who notes that licensure of traditionally prepared teachers has a direct impact on increasing student achievement; by the National Assessment of Educational Progress (2018) who reports scores in mathematics are greater for students taught by fully certified teachers. The latter was borne out by the results of the current study with significance for traditionally prepared teachers teaching 5th and 11th grade math achievement favoring traditionally prepared subjects (Darling-Hammond, 2002).

Teacher licensure and certification scores impact student achievement but experience is a constant, pervading all content areas and levels (Clotfelter, Ladd, & Vigdor 2010). The latter was also supported in the current study by the significance found for math and English scores at the 5th and 11th grades moderated by the experience variable and the progression of significance from 0 to 10 years.

Again, the impact mentioned in the above cited studies support the current findings in regard to the significance of these variables on student achievement. Laczko-Kerr, and Berliner (2002) found that Teach for America students did not perform much differently than did students taught by noncertified teachers and that students taught by certified teachers performed greater than students taught by noncertified teachers. Their data supports the findings of this study regarding certification status.

In a study of Teach for America in Arizona schools, students taught by noncertified teachers did significantly poorer on math, reading and English tests (Laczko-Kerr & Berliner, 2002). A meta-analysis study comparing alternative certification pathways and traditionally prepared teachers found significant but small differences favoring alternatively prepared teachers' effect on student achievement (Allen, 2003). Mean achievement differences of students

taught by traditionally prepared teachers was .03 of a standard deviation less than students taught by alternatively prepared teachers. The current study showed significance for traditionally prepared teachers at both grade levels and content areas.

As reported by the American Association of Colleges for Teacher Education (2012), a survey was conducted with 224 first year alternatively prepared teachers and 577 traditionally prepared teachers about how they felt about their preparation. About half of the alternatively prepared teachers felt prepared compared to 80% of the traditionally prepared teachers. More than half of the alternatively prepared teachers said they had too little time working with a licensed classroom teacher before their assignment, compared to 20% for traditionally prepared teachers. Ninety-four (94) percent of traditionally prepared teachers expressed confidence that their students were learning and responded to their teaching, compared to 74% for alternatively prepared teachers. One of the experience factors in the current study was 0-3 years of experience but was not sorted out for first-year teachers. However, that result did show significance for the 0-3 factor favoring traditionally prepared teachers compared to alternatively prepared teachers.

Although it is not known in this study about the various pathways taken by the alternatively prepared teachers in the current sample, the results were consistent that students who were taught by alternatively prepared teachers scored significantly lower on the year-end WV summative assessments in math and English compared to traditionally prepared teachers.

Research has confirmed that teachers with national board certification have a special quality and can impact student achievement (Cowen & Goldhaber, 2015). Such impact would not be unexpected because of the rigorous standards, intense training and self-preparation engaged in by these teachers. The current study did not identify the numbers of board certified alternatively or traditionally prepared teachers; however that effect was likely present to a degree

affecting student achievement because of the whole of hundreds of traditionally prepared teachers in the sample who moderated significant gains in student achievement. It might have been expected that students in 5th grade English and mathematics would have lesser content achievement (test scores) given that it is unlikely that many of their teachers would have earned a master's degree in a core content discipline or completed a well-defined specialization in math and English, i.e., 15 semester hours or more in these fields. However, the current results bear otherwise; students taught by traditionally prepared teachers significantly scored greater on summative tests compared to their peers taught by alternatively prepared teachers.

Blank (2007) offers that 10% of the gain for student achievement is attributed to teacher experience but plateaus at 5 years of experience for elementary teachers and between 5- 9 years for secondary. The author also added that content knowledge becomes more important and complicated as grade levels increase. Rice (2003) contends that teacher experience is more influential at the secondary level. Current results were not necessarily plateaus per se; however, these showed overall that as teacher experience progressed student achievement increased, with the greatest impact at 10+ years for both 5th and 11th grade levels in math and English.
In contrast, current results differed from Wiswall (2013) who believed that experience is not a major factor for student achievement. The latter indicated that 5th grade teachers did not produce better results after the first few years of teaching. That outcome was not supported by the results of the current study, which showed student achievement progressed as teacher experience progressed.

Huang and Moon (2009) indicated that experience was not a constant across all grade levels and content areas. While the current study only included two grade levels and two subject areas, its results clearly showed experience to be a major factor. The latter is also supported by

the results of a longitudinal study by Ladd and Sorensen (2014), who reported long-term effects of experience on student achievement, student behavior and attendance. It seems logical that teacher experience leads to teacher "know how" which indirectly or directly has a positive effect on student achievement—and on test scores in the case of the current results.

Various researchers point out that it may not be in the best interests of students and society to give a new alternatively prepared teacher the immediate responsibility for day-to-day classroom managerial and instructional functions prior to the completion of a teacher preparation program. Noncertified teachers who have full responsibility can easily struggle more so than fully trained teachers (Darling-Hammond, 2002). Additionally, underprepared teachers in alternative preparation programs often tend to be employed in districts and schools with greater percentages of academically at-risk students (Alexander & Fuller, 2011). Others offer that it is too simplistic to assume that having a subject matter degree and some type of professional support and mentoring will yield effective instruction or to expect that teaching abilities can be quickly developed on the job (Ovando & Trube, 2000). Overall, alternatively prepared teachers in the current study did not significantly impact student achievement when compared to traditionally prepared teachers.

What is the effect of teacher quality on student learning and achievement? Teacher quality is an elusive concept and a common meaning has challenged professional educators and policy makers. Its variables and characteristics vary from state to state and district to district. It may be commonly understood that teacher quality is a dynamic concept that includes one's instructional and managerial competence, depth of content knowledge, verbal ability and articulation, adaptiveness, and personal countenance. However, the standard proxy to assess teacher quality most commonly includes completion of degrees, passing the state's competency

test in the discipline and being certified in a core subject. Experience does not seem to be a factor. Add to this issue, the former goal (requirement) of state agencies and the federal government for teachers to be highly qualified (HQT).

Given that a great majority of teachers nationally (and in WV) have achieved HQT status, there remains no substantial evidence linking a change in quality to student achievement (Rothman, 2009). Goe (2007) and Betts, Zau, & Rice (2003) offer that full certification in the content major has a greater impact on student achievement. HQT teachers in the current study did have a significant impact on student achievement compared to their non-HQT peers. But again, the proxy is summative test scores. Yet, many offer that standardized test scores are the best data for measuring student achievement and attributing teacher quality. Notwithstanding the known downsides, such tests when carefully crafted are secure, comprehensive, and comparable across schools and grade levels and have usually passed reliability and validity standards.

Teacher quality can be an elusive concept but it is also important to examine the characteristics associated with non-HQT teachers. What are its common characteristics or specifically what do these teachers lack? Once that is identified, particular attention can be given to addressing these needs programmatically, i.e., professional development, support and mentoring, more frequent supervisory evaluation and feedback,

State level educators and policy makers can apply the results of the current study to more fully evaluate and understand the connection of summative test scores to teacher licensure, teacher experience, teacher preparation and teacher quality. The current study can highlight the need to reassess policies regarding licensure and teacher certification and be used as a foundation for examining the pathways that lead to teachers becoming certified.

Consideration should be given to the kinds of teaching placements and assignments given to noncertified teachers, as questioned by Boyd, Goldhaber, Lankford, & Wyckoff (2007). Are the least prepared teachers assigned to teach the neediest and most academically challenging youngsters? Notwithstanding the pressure of teacher shortages, some careful attention needs to be given to these placements/assignments.

There is a need to examine the reasons behind the exceptionally large numbers of students whose summative test scores placed them in the categories of "standard not met" and "standard almost met." For example, 54% of 11th grade students did not meet the math standard, along with 27% who "almost" met. A similar, but lesser effect occurred for 5th grade math achievement: 39% did not meet the standard along with 30% who almost did. For English 11 students, 28% did not meet and 27% almost met. For English 5, 19% met the standard and 16% almost did. Eighty-one percent (81) of math 11 students did not meet or almost met the achievement expectations in mathematics and 55% did the same for English 11. Notwithstanding the significance for the traditionally prepared teacher variable, 11th grade teachers have students with a high rate for not meeting standards, particularly for mathematics.

Of course, it is the scaled summative test scores of these students that result in these placements. If summative test scores are going to be the proxy for student achievement, then it needs to be known if these scores are "optimal"—that students are motivated to do well on these tests. King (2017) in a study of WV high school students found that 10th graders gave variant responses to the importance attached to and the effort expended when taking their year-end summative tests. A conclusion was that such tests do not have a high stake for students—no consequence per se. Additionally, it is important to determine what kind of preparation occurs, how often and if it is a systematic approach. It is important to note that West Virginia Math

scores for all students, as shown in results from the 2017 summative assessment, indicate that only 34% of all students are proficient in mathematics (WVDE, 2017). Therefore, there are additional variables that have an impact on mathematics achievement beyond the ones presented in this study.

Recommendations for Further Study

Examine differences in the data comparing teachers trained outside of WV to those trained in WV. This could also include those who were alternatively prepared and traditionally prepared.

The research provided a snapshot of what existed in 2015-2016—replicate the study to determine if similar effects currently hold.

Build on a previous study (King, 2017) examining the level of interest, motivation and effort students bring to the year-end, summative testing context.

Determine what percentage of math and English courses at the 5^{th} and 11^{th} grade levels were taught by non-HQT and HQT teachers.

Determine the proportion of 5^{th} grade teachers who completed a second teaching specialization in a content discipline or a well-defined specialization.

Compare teachers and students in WV counties considered economically depressed or in low social and economic circumstances.

Employ a predictive data analysis model for the existing variables and data which can assess the data as separate predictors and also account for combined effects.

Add a qualitative component—personal interviews of selected math and English teachers to obtain in-depth knowledge about the teacher quality variable.

An annual report of the U.S. Department of Education (AACTE, 2012) states that the vast majority of alternative preparation programs are affiliated with an institution of higher education (92%). Propose similar arrangements (affiliation with higher education) with state and locally sponsored alternative preparation programs.

Investigate the effects on the study variables with samples of first year alternatively prepared and traditionally prepared to determine what effects/differences may be operating at the outset.

Summary

The results of the current study have shown the differential effects of types of teacher preparation pathways. The results were supported by many related studies in the existing research literature, while at the same time, there were many studies that countered the results in favor of alternative preparation programs. This seems to have been a trend over the past 30 years or so—variations and inconclusive evidence. Probably the "best" way to educate professional public-school teachers begs the question. There are very likely several "best" ways, depending on a number and kind of contextual variables and circumstances. What may work in West Virginia may not generalize outside of the state.

The various research studies published in the existing research literature during the evolution of alternative preparation programs and traditional preparation programs effects have a mix of research methodology, variables, and outcome measures. Also, these have largely been conducted at the local level, even though sponsored by federal, state and private resources. The current study was only one of several projects conducted on a larger scale at the state level. Also, many of the studies in the literature were designed with a limited focus: to compare the alternative preparation and traditional teacher preparation variable against a single outcome. The

current study designed in four variables and applied a summative content assessment outcome to two grade levels and two content disciplines. In effect it had a multivariate focus. The result is that one can perceive the separate and combined effects of the study's variables on student achievement.

As noted, the research literature has mixed results on the effects of alternative preparation and traditional preparation pathways. An importance of the current investigation is its contribution to the national database on alternative preparation and traditional preparation and to understand the "best" ways to prepare classroom teachers who can have the "best" impact on their students' learning and achievement.

REFERENCES

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics*, *25*(1), 95–135.

Alexander, C. & Fuller, E. (2011). Does teacher certification matter? Teacher certification and middle school mathematics achievement in Texas. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Allen, M. (2003). *Eight questions on teacher preparation: What does the research say?* Education Commission of The States. Denver, CO.

American Association of Colleges for Teacher Education (2012). *Where we stand: Alternative vs. traditional teacher preparation.* Retrieved from AACTE website: https://secure.aacte.org/apps/rl/res_get.php?fid=481&ref=rl

Betts, P. & Frost, L. (2000). Subject knowledge and teacher preparation. *Education Canada*, (40)1, 2.

Betts, J. Zau, A., & Rice, L. (2003). *Determinants of Student Achievement: New Evidence from San Diego.* San Francisco: Public Policy Institute of California.

Behrstock-Sherratt, E. & Potemski, A. (2013). Performance-based compensation: linking performance to teacher salaries. Ask the team. Center on Great Teachers and Leaders. Retrieved from ERIC database. **(**ED555667)

Blank, L. (2007). Teach for America setting sights on Pre-K. *Education Week*, *26*(33), 1-16.

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *Future Of Children*, *17*(1), 45-68.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008). *Teacher preparation and student achievement.* National Bureau of Economic Research. (Working Paper 14314). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w14314.pdf

Brinkman, B. (2014). The impact of teacher preparation program on academic achievement of eighth grade students in a south Texas school district. (Doctoral Dissertation). Texas A&M University-Corpus Christi Corpus Christi, Texas. Retrieved from https://tamucc-ir.tdl.org/tamucc-ir/bitstream/handle/1969.6/560/Barbara%20Brinkman.pdf?sequence=1

Budding, R., & Zamarro, G. (2009). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics*, *66*(2), 103-115.

Carr, M. (2006). *The determinants of student achievement in Ohio's public schools.* (Policy Report). Columbus, OH: Buckeye Institute for Public Policy Solutions.

Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Human Resources Summer*, *45*(3), 655-681.

Clotfelter, C., Ladd, H.F., & Vigdor, J.L. (2007). *Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects.* Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w13617

Clotfelter, C., Ladd, H.F., & Vigdor, J.L. (2007a). How and why do teacher credentials matter for student achievement? CALDER Working Paper 2. Washington, DC: The Urban Institute.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Cohen, J. & Goldhaber, D. (2015). Observations on different means of evaluating teacher performance. In J.A. Grissom & P. Youngs, eds.: *Making the Most of Multiple Measures: The impacts and Challenges of Implementing Rigorous Teacher Evaluation Systems.* New York: Teachers College Press.

Cowan, J., & Goldhaber, D. (2015). National board certification and teacher standards. Evaluating and empowering teachers. Retrieved from National Board for Professional Teaching Standards website: https://www.nbpts.org/research

Cronigner, R.G., Rice, J.K., Rathun, A., & Nishio, M. (2003). *Teacher qualifications and first grade achievement: A multilevel analysis.* College Park, MD: Center for Education Policy and Leadership.

Darling-Hammond, L. (2007). Research and rhetoric on teacher certification. Retrieved from Researchgate website: https://www.researchgate.net/publication/49610113_Research_and_Rhetoric_on_Teacher_Certification

Darling-Hammond, L. (2000). Reforming teacher preparation and licensing: debating the evidence. *Teachers College Record, 102*(1), 28-56.

Darling-Hammond, L. (2000a). How teacher education matters. *Journal of Teacher Education, 51*(3), 166-173.

Darling-Hammond, L., & Sykes, G. (2003). Wanted: A national teacher supply policy for education: The right way to meet the "Highly Qualified Teacher" challenge. *Educational Policy Analysis Archives, 11*(33).

Darling-Hammond, L. (2002). Research and rhetoric on teacher certification: A response to teacher certification reconsidered. *Educational Policy Analysis Archives, 10*(36).

Decker, P., Mayer, D., & Glazerman, S. (2004). *The effects of Teach for America on students: Findings from a national evaluation*. Princeton, NJ: Mathematics Policy Research Inc.

Educational Research Newsletter (2003). *Comparing alternative and traditional teacher preparation.* Retrieved from https://www.ernweb.com/educational-research-articles/comparing-alternative-and-traditional-teacher-preparation/

FindLaw. No Child Left Behind Act and teacher accountability (2009). Retrieved from FindLaw website: http://files.findlaw.com/pdf/education/education.findlaw.com_curriculum-standards-school-funding_no-child-left-behind-act-and-teacher-accountability.pdf

Fowler, R.C. (2003). The Massachusetts signing bonus program for new teachers; A model for teacher preparation worth copying? *Education Policy Analysis*, *11*(13).

Fuller, E., & Alexander, C. (2003). *Teachers in Texas from different routes to certification: Who are they, where do they teach, and how long do they teach?* Paper presented at the American Educational Research Association Annual Meeting, Chicago, IL.

Gawthrop, J. (2014). Measuring student achievement: A study of standardized testing & its effect on student learning. Retrieved form Jessup website: http://my.jessup.edu/publicpolicy/wp-content/uploads/sites/39/2014/04/Gawthrop_Jeremiah_Final.pdf

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 & Urban Society*, *40*(1), 91-117.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality.

Goldhaber, D.D., & Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Education Evaluation and Policy Analysis*, *22*, 129-145.

Goldhaber, D.D. & Brewer, D.J. (1999). *Teacher licensing and student achievement*. In C. Finn and M. Kanstoroom, eds.: *Better Teachers, Better Schools.* Washington, D.C. Thomas B. Fordham Institute. Retrieved from: https://www.stcloudstate.edu/tpi/initiative/documents/recuitment/Teacher%20Licensing%20and%20Student%20Achievement.pdf

Harris, D.N., & Sass, T. (2007). *Teacher training, teacher quality, and student achievement* (Working Paper No.3). Washington, DC: National Center for Analysis of Longitudinal Data in Education Research.

Harris, D.N. & Sass, T. (2013) Licensure and worker quality: A comparison of alternative routes to teacher certification. *Journal of Public Economics*, *95*(7-8), 798-812.

Holloway, Ellen (2007). Do qualifications imply quality? The relationship between "highly qualified teachers" and student achievement in reading and math in elementary school. Retrieved from https://repository.library.georgetown.edu/bitstream/handle/10822/555861/etd_eth2.pdf;sequence=1

Huang, F., and Moon, T. (2009). Is Experience the Best Teacher? A Multilevel Analysis of Teacher Characteristics and Student Achievement in Low Performing Schools. *Educational Assessment, Evaluation and Accountability*, *21*(3), 209-234.

Ingersoll, R. & Strong. W. (2015). *The impact of induction and mentoring programs for beginning teachers: A critical review of the research in California*. Graduate School of Education Berkeley: University of California.

Ingersoll, R., Merrill, L., & May, H. (2014). *What are the effects of teacher education and preparation on beginning teacher attrition?* Research Report (#RR-82). Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.

Kane, T.J., Rockoff, J.E., & Staiger, D.O. (2008). What does certification tell us about teacher effectiveness? Evidence From New York City. *Economics of Education Review*, *27*(6), 615-631.

King, R. (2017). *A comparison of self-perceived test taking motivation on large scale high and low stakes tests*. (Unpublished Doctoral Dissertation). Marshall University, Huntington, WV.

Koedel, C., and Betts, J. (2007). *Re-examining the role of teacher quality in the educational production function.* Working Paper 2007-03. National Center on Performance Incentives. Retrieved from https://economics.missouri.edu/working-papers/2007/wp0708_koedel.pdf

Laczko-Kerr, I. & Berliner D.C. (2002). The effectiveness of Teach for America and other undercertified teachers on student achievement: A case for harmful public policy. *Educational Policy Analysis Archives*, *10*(37).

Ladd, H.F., & Sorensen, L. (2014). *Returns to teacher experience: Student achievement and motivation in middle school.* National Center for analysis of Longitudinal Data in Education Research. Retrieved from Caldercenter website: http://www.caldercenter.org/sites/default/files/WP%20112%20Update_0.pdf

Mader, J. (2013). *Alternative routes to teaching become more popular despite lack of evidence*. The Hechinger Report. Retrieved from Hechinger Report website: http://hechingerreport.org/content/alternative-routes-to-teaching-become-more-popular-despite-lack-of-evidence_12059/

National Assessment of Educational Progress (2018). *2017 NAEP mathematics & reading assessments.* Retrieved from the Nation's Report Card website: https://www.nationsreportcard.gov/reading_math_2017_highlights/

National Commission on Teaching & America's Future (2016). *What matters now: A New compact for teaching and learning.* Arlington, VA. Retrieved from National Commission on Teaching & America's Future website: https://nctaf.org/wp-content/uploads/2016/09/NCTAF_What-Matters-Now_The-Evidence-Base_hyperlinked.pdf

National Council on Teacher Quality (2013). *Teacher preparation program student performance data models: Six core design principles*. Washington D.C. Retrieved from the National Council on Teacher Quality website: https://www.nctq.org/dmsView/Teacher_Preparation_Program_Student_Performance_Data_Models_NCTQ_Report.

Nyankori, R.J. (2005). *Does certification of elementary school teachers matter? The effects of certification status on instructional practices and on the mathematics and reading achievement of first grade public school students*. Dissertation submitted to the University of Maryland, College Park.

Ovando, M., & Trube, M. (2000). Capacity building of beginning teachers from alternative certification programs: Implications for instructional leadership. *Journal of School Leadership, 10*, 346-366.

Peterson, P., & Nadler, D. (Winter 2009). What happens when states have genuine alternative certification? *Education Next*, *9*(1), 70-74.

Program for International Student Assessment (2015). *PISA 2015 results.* Washington, DC: National Center for Education Statistics.

Redding, C., & Smith, T. (2016). Easy in, easy out: Are alternatively certified teachers turning over at increased rates? *American Educational Research Journal, 53*(4), 1086-1125.

Rice, J.K. (2010). *The impact of teacher experience: examining the evidence and policy implications*. Retrieved from Urban Institute website: http://www.urban.org/sites/default/files/publication/33321/1001455-The-Impact-of-Teacher-Experience.PDF

Rice, J.K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.

Riordan, J. (2009). *Do teacher qualifications matter? A longitudinal study investigating the cumulative effect of NCLB teacher qualifications on the achievement of elementary school children.* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.

Robinson, H. (2011). Highly qualified teacher status and the reading achievement of students with disabilities. *American Secondary Education*, *39*(3), 42-66.

Rosales, J. (2014). *Pay based on test scores? What educators need to know about linking teacher pay to student achievement*. Retrieved from NEA website: http://www.nea.org/home/36780.htm

Rothman, Robert (2009). *Does Highly Qualified Mean Highly Effective?* Center for Public Education. Retrieved from Center for Public Education website: http://www.centerforpubliceducation.org/Main-Menu/Staffingstudents/How-good-are-your-teachers-Trying-to-define-teacher-quality/Does-highly-qualified-mean-highly-effective.html

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *Quarterly Journal of Economics*, *125*(1), 175–214.

Ryan, J.E. (2004). The perverse incentives of the No Child Left Behind act. *New York University Law Review*, *79*(932).

Sanders, W.L. (1998). Value added assessment. *School Administrator 11*(55), 24-27.

Sanders, W.L. & J.C. Rivers (1996). *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.

Sawchuk, S. (2011). *What studies say about teacher effectiveness.* The National Education Writers Association. Retrieved from EWA website: http://www.ewa.org/sites/main/files/file-attachments/teachereffectiveness.final_.pdf

Sass, T.R. (2013). *Licensure and worker quality: A comparison of alternative routes to teaching. SSRN Electronic Journal*, *58*(1).

Schleicher, A. (2013). *Lessons from PISA outcomes*. OECD Observer, No 297 Q4 2013.

Smith, J.B., Lee, V.E., & Newmann, F.M. (2001). *Instruction and achievement in Chicago elementary schools: Improving Chicago schools.* Chicago: Consortium on Chicago School Research. Retrieved from http://ccsr.uchicago.edu/sites/default/files/publications/p0f01.pdf

Snyder, J., & Bristol, T.J. (2015). Professional accountability for improving life, college, and career readiness. *Education Policy Analysis Archives*, *23*(16). Retrieved from http://files.eric.ed.gov/fulltext/EJ1070474.pdf

Springer, M., Ballou, D., Hamilton, L., Le, V., Lockwood, J., McCaffrey, D., Pepper, M., and Stecher, B. (2010). *Teacher pay for performance: Experimental evidence from the project on incentives in teaching.* National Center for Performance Incentives. Retrieved from

RAND website: https://www.rand.org/content/dam/rand/pubs/reprints/2010/RAND_RP1416.pdf

Suell, J. & Piotrowski, C. (2006). Efficacy of alternative teacher certification programs: a study of the Florida model. *Education*, *127*(2), 310-315.

Tucker, P, & Stronge, J. (2005). *Linking Teacher Evaluation and Student Learning*. Alexandria, VA: Association for Supervision and Curriculum Development.

U.S. Department of Education. (2004). *Innovations in education: Alternative routes to teacher certification.* Office of Innovation and Improvement. Retrieved from https://www2.ed.gov/admins/tchrqual/recruit/altroutes/report.pdf

U.S. Department of Education (2014). *New No Child Left Behind flexibility: Highly qualified teachers*. Fact Sheet. Retrieved from https://www2.ed.gov/nclb/methods/teachers/hqtflexibility.html

U.S. Department of Education. (2015). *Every Student Succeeds Act*. Washington, DC. Retrieved from https://www.ed.gov/essa

U.S. Department of Education. (2015a). *Title II reports*. Retrieved from https://title2.ed.gov/Public/Home.aspx

Walsh, K., Joseph, N., & Lewis, A. (2016). *Within our grasp: Achieving higher admissions standards in teacher prep*. 2016 State Teacher Policy Yearbook Report Series. Retrieved from National Council on Teacher Quality website: https://www.nctq.org/dmsView/Admissions_Yearbook_Report

Wenglinsky, H. (2000). *How teaching matters. Bringing the classroom back into discussions of teacher quality.* A Policy Information Center Report. Retrieved from ETS website: https://www.ets.org/Media/Research/pdf/PICTEAMAT.pdf

West Virginia Department of Education (2017). *Comparison of West Virginia statewide assessment results*. Retrieved from WV Department of Education website: http://static.k12.wv.us/tt/2017/2017comparisonstatewideresults.pdf

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, *100*(61), 7.

APPENDIX A: APPROVAL LETTER

MARSHALL UNIVERSITY
www.marshall.edu

Office of Research Integrity

February 28, 2018

Robert Hagerman
81 Derek Drive
Hurricane, WV 25526

Dear Mr. Hagerman:

This letter is in response to the submitted thesis abstract entitled *"Impact of Educator Preparation Routes and Teacher Characteristics on Student Achievement."* After assessing the abstract, it has been deemed not to be human subject research and therefore exempt from oversight of the Marshall University Institutional Review Board (IRB). The Code of Federal Regulations (45CFR46) has set forth the criteria utilized in making this determination. Since the information in this study does not involve human subjects as defined in the above referenced instruction, it is not considered human subject research. If there are any changes to the abstract you provided then you would need to resubmit that information to the Office of Research Integrity for review and a determination.

I appreciate your willingness to submit the abstract for determination. Please feel free to contact the Office of Research Integrity if you have any questions regarding future protocols that may require IRB review.

Sincerely,

Bruce F. Day, ThD, CIP
Director

APPENDIX B: SCALE SCORE RANGE

West Virginia Summative Assessment Reported Scale Scores Range – *Level 1, Level 2, Level 3, and Level 4.*

ELA/Literacy Reported Scale Scores				
Grade	**Level 1**	**Level 2**	**Level 3**	**Level 4**
5	2056 - 2441	2442 -2501	2502 - 2581	2582 - 2916
11	2102 - 2492	2493 - 2582	2583 - 2681	2682 - 3032

Mathematics Reported Scale Scores				
Grade	**Level 1**	**Level 2**	**Level 3**	**Level 4**
5	2095 - 2454	2455 - 2527	2528 - 2578	2579 - 2891
11	2118 - 2542	2543 - 2627	2628 - 2717	2718 - 3085

APPENDIX C: VITA

Robert Hagerman
Charleston, WV
srhagerman@suddenlink.net

Education

Marshall University (2018)
Ed.D. Curriculum and Instruction

Marshall University (1996)
Master's in Secondary Education and Educational Computing

West Virginia State University, Magna Cum Laude (1995)
Bachelor of Science, Secondary Education

Summary of Qualifications

Education professional with 18 years of experience in teaching K-12 students and adult learners. Extensive experience in building partnerships with a variety of public and private agencies, organizations, and institutions of higher education at the local and national level. Extensive knowledge of federal, state, and local policies and regulations related to education at the P-12 and higher education levels. Expertise in financial management, budget creation, monitoring, and reporting. Over ten years of experience in providing leadership to a large constituency including policy makers in the areas of educational policy with primary responsibilities for approval and quality control of education programs. Experienced in drafting, reviewing, and submitting successful grant applications to public, private, and federal sources including expertise in education research and development.

Skills and Expertise

- Public policy
- Policy writing
- Program accreditation and review
- Management and supervision
- Grant writing
- Program coordination and supervision
- Staff development
- Workshops and seminars
- Budget development, implementation, and maintenance

Professional Experience

Executive Director, Office of Certification and Professional Preparation, West Virginia Department of Education (2016-Present)

Develop, implement, and provide leadership on issues of policy, W. Va. State Code, and federal guidelines for matters related to educator preparation accreditation and program review of traditional and alternative routes, educator quality including teacher retention and recruitment, testing, professional development, and federal programs. Facilitate the alignment of educator preparation programs and the quality of its candidates with state needs and priorities.

Assistant Director, *Office of Certification and Professional Preparation*, West Virginia Department of Education

Teacher Quality Coordinator, *Office of Certification and Professional Preparation*, West Virginia Department of Education

English Language Learner/ESL Education Specialist and Partner Liaison

Appalachia Regional Comprehensive Center at Edvantia, Charleston, West Virginia

Program/services administrator; partnership-building, coordination and supervision of work among a large network of regional and national partner organizations; ESL consultant for program serving five state departments of education; workshop presentations and technical assistant provider to build capacity in matters related to education at the P-20 level. Worked with and managed large budgets for projects involving individual schools, districts, consortia, and state agencies. Conducted Instructional Learning Appraisals at the school and district levels across states within the organization's multi-state territory.

Technology Specialist/Coordinator of User Services, ERIC Clearinghouse on Rural Education and Small Schools at AEL, Charleston, West Virginia

Areas of responsibility included dissemination of educational research information, report and publication writer, workshop presentations, virtual digital reference sessions, in the areas of rural education, Mexican-American education, migrant education, American-Indian education, and small schools. Part of a national and international information dissemination network.

Spanish teacher, West Virginia Virtual School (West Virginia Department of Education). Teacher for rural counties in West Virginia

Spanish Teacher/Computer Teacher/Technology Coordinator, Fayette and Kanawha County Schools, West Virginia

Adjunct Professor, West Virginia State University, Institute, WV