



Seattle Pacific University
LIBRARY
Discover, Create, Share

Seattle Pacific University
Digital Commons @ SPU

Education Dissertations

Education, School of

Spring April 19th, 2018

Differences in Performance on a Commercially Available Competency-Based Exam between Traditionally and Alternatively Prepared Teacher Candidates: A Causal Comparative Study

Lara Denise Cole

Follow this and additional works at: https://digitalcommons.spu.edu/soe_etd



Part of the [Teacher Education and Professional Development Commons](#)

Recommended Citation

Cole, Lara Denise, "Differences in Performance on a Commercially Available Competency-Based Exam between Traditionally and Alternatively Prepared Teacher Candidates: A Causal Comparative Study" (2018). *Education Dissertations*. 36.
https://digitalcommons.spu.edu/soe_etd/36

This Dissertation is brought to you for free and open access by the Education, School of at Digital Commons @ SPU. It has been accepted for inclusion in Education Dissertations by an authorized administrator of Digital Commons @ SPU.

Differences in Performance on a Commercially Available Competency-Based Exam
between Traditionally and Alternatively Prepared Teacher Candidates:
A Causal Comparative Study

Lara D. Cole

Seattle Pacific University

Differences in Performance on a Commercially Available Competency-Based Exam
between Traditionally and Alternatively Prepared Teacher Candidates:

A Causal Comparative Study

by

Lara D. Cole

Dissertation

Seattle Pacific University

Defense:

Approved by _____
(Dr. Arthur Ellis, Ed.D., Chairman of the Dissertation Committee)

(Dr. Nyaradzo Mvududu, Ed.D., Committee Member)

(Dr. William Prenevost, Ed.D., Committee Member)

Program Authorized to Offer Degree: School of Education

Date _____

(Dr. Rick Eigenbrood, Ph.D., Dean, School of Education)

Copyright Page

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at Seattle Pacific University, I agree that the library shall make its copies freely available for inspection. I further agree that extensive copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106, to whom the author has granted the "right to reproduce and sell (a) copies of the manuscript in microfilm and/or (b) printed copies of the manuscript from microfilm."

Signature _____

Date _____

TABLE OF CONTENTS

	<i>Page</i>
List of Tables	ii
Acknowledgments	iii
ABSTRACT.....	1
Chapter One Introduction	2
Chapter Two Review of Literature	11
Chapter Two Summary	35
Chapter Three Methodology	40
Chapter Four Results	53
Chapter Four Summary	61
Chapter Five Discussion	62
References	74

LIST OF TABLES

	<i>Page</i>
Figure 1 Ethnicity of Students Enrolled in the School of Education	44
Table 1 Number of edTPA Content Area Tests Administered by Year	46
Table 2 Contingency Table Overall edTPA Score	56
Table 3 Descriptive Statistics.....	59

ACKNOWLEDGMENTS

I would like to thank Dr. Arthur Ellis, Dr. Nyaradzo Mvududu, and Dr. William Prevenost for their guidance in this endeavor. Many others (formal and informal mentors) were also invaluable in this process.

Bob Butler, Barbara Butler, William Hewitt, and Danika Cole provided support and encouragement throughout the doctoral program and witnessed the process on a daily basis. They exemplify tenacity and patience; both qualities necessary for a doctoral student.

Finally, I would like to thank Janna Pope, an outstanding professional and advocate for children, and Kenny. May they continue to serve as markers to guide my future endeavors.

ABSTRACT

Alternative route teacher education programs have increased to address personnel shortages in key areas such as special education, mathematics, and science. Alternative route programs may serve as a means to address such shortages, but require evaluation to ensure that candidates emerging from these programs demonstrate skills commensurate with those that complete traditional teacher education programs. Although studies examining aspects of alternative route programs exist, few studies comparing alternative route programs with traditional programs have been conducted. Differences in definitions used to describe alternative route programs compound efforts to compare programs. Nevertheless, research reviewed compared outcomes for alternative and traditional programs on candidate satisfaction, academic achievement, multicultural awareness, retention, and performance on competency-based assessments. Results on these measures were mixed.

The purpose of this causal comparative study was to compare performance on the edTPA (Teacher Performance Assessment) using longitudinal data (2013-2017) from candidates ($N = 565$) that attended a mid-sized urban university that offered both traditional and alternative route programs. The non-parametric analysis revealed no differences in expected and actual edTPA performance between groups. The parametric analysis revealed a statistically significant omnibus effect based on program type. Post hoc testing revealed differences in Instructional Average (IA) and Assessment Average (AA) although variance accounted for was small. Recommendations for practitioners are discussed.

Chapter One

Introduction

Context

Mencken (n.d.) stated, “For every complex problem, there is an answer that is clear, simple, and wrong.” Although Mencken may not have been speaking directly to issues of school governance, teacher quality, and labor shortages, he very well could have been. The purpose of this study is to examine an aspect of public schooling in the United States; specifically, teacher shortages and attempts to address the problem through the creation of new teacher preparation programs, and how (if) we have evaluated the results of these programs. It is unlikely that a clear and straightforward solution will emerge. This present study represents an attempt to provide comparisons of efficacy between traditional and alternative routes to certification programs, an area that has received little attention (Aragon, 2016; Hussar & Bailey, 2016).

Teacher shortages. Claims of teacher shortages belie current data, at least upon initial analysis. Hussar and Bailey (2016) reported that total elementary and secondary enrollment is projected to increase by 4% by 2023. The number of elementary and public school teachers is projected to rise by 8% by 2023 (2016). An increase in both student enrollment and teachers seems promising. Unfortunately, the increases in student enrollment are not consistent across geographic regions in the United States; 34 states are predicted to have a 5% increase in enrollment by 2023, while 16 states are predicted to have a decrease in enrollment of 5% by 2023. Increasing (shifting) student enrollment and overall increases in the number of teachers projected to enter the field do not illuminate an underlying problem: teacher shortages in key areas. The U.S. Department

of Education (ED, 2017) compiled a list of teacher shortage areas nationally and by state. Overall results indicated that shortages exist in special education, science, and mathematics. Teacher shortages in other content areas also existed, yet shortages in special education, science, and mathematics were consistent in nearly every individual state (2017).

Response to teacher shortages. National, state, and local efforts have attempted to ameliorate the impacts of teacher shortages. Students qualifying for federal financial aid may be eligible for partial or full loan forgiveness if they agree to serve in an area identified as a key shortage area (ED, 2017). Many states have developed alternative routes to certification programs (in conjunction with teacher preparation programs and school districts). The ED (2004) has hailed these programs as innovations. Estimates indicate that at least 20% of new teachers entering the teaching profession will do so via an alternative route program (Woods, 2016).

Evaluation of teacher education programs: Early and ongoing efforts. The Common School Movement of the 19th century represented an increase in the number of schools in our country (Boers, 2007). The state of Massachusetts alone established 1,500 new public schools between 1840 and 1860 and this expansion continued westward with the population (Ramsey, 2014). Efforts to supervise the provision of services in these schools increased during this time with the development of county and city superintendents who were responsible for both schools and curriculum. As attendance increased, superintendents required additional support, and additional layers of bureaucracy emerged. Horace Mann is the oft-cited “father of the Common School Movement” and made innumerable efforts to ensure that students received quality

services (Boers, 2007). Programs and practices established by Mann as he served as the first Secretary of Education in Massachusetts spread throughout the country (2007). In addition to on-site supervision of schools, Mann established the “Common School Journal,” which was a professional journal of sorts designed to help disseminate his guiding principles for the education of pupils. He also advocated for increased professional regard for teaching as a profession and established “normal schools” to serve as formalized preparation programs for future teachers (2007).

Many of the regulatory structures Mann developed remain with us today, although the regulatory milieu public schools face is exponentially more complex as is the range of students served. Federal versus local control has remained a theme following the Common School Movement. Early legislative efforts to improve schools and ensure teacher quality reflected these tensions (Chopin, 2013). In 1867, the ED was established (2013). The National Defense Education Act (Public Law 85-864) passed in 1958 provided assistance and guidance to state and local school systems to strengthen public school instruction and improve instruction offered by institutions of higher education (2013). The Elementary and Secondary Education Act (ESEA) of 1965 provided grants aimed explicitly toward increasing quality in teacher preparation programs (2013). No Child Left Behind (NCLB, passed in 2002 with bipartisan support) is the most recent reauthorization of the ESEA of 1969 (Chopin, 2013). NCLB included provisions specifically designed to enhance student achievement and ensure school quality. Measures to rank schools, award districts/schools with recognition or penalties based upon performance, and criteria for educators to be “highly qualified” were included (2003). While hailed as a success for increasing attention to the achievement of

subgroups, many criticized one of NCLB's provisions: the requirement that all students reach proficiency on a high stakes exam by the year 2014 (2003). Revisions to NCLB are ongoing, and the nomenclature will change as political appointees and elected officials wrestle with various educational complexities.

In 2009 Race to the Top (a federal grant incentive program) followed NCLB (ED, 2015). The primary objectives of this voluntary grant application were to ensure that states who applied for funds legislated teacher/principal evaluation systems that included student growth, adopted the Common Core Standards, participated in the Smarter Balanced Assessment System, and built infrastructure to evaluate and disseminate student achievement data. At least 45 states submitted applications to the program; nineteen states received an award. In excess of 700 million dollars was allotted toward this program. The ED (2015) cited many positive results of the program. Unfortunately, Arne Duncan, who served as U.S. Secretary of Education when his department issued the final report on the program, was reluctant to cite increased student achievement (a critical measure of a program's success) as an outcome stating, "Race to the Top's success ultimately must be measured by its long-term impact on student learning. Because simultaneous change in multiple systems takes time, it is too early to make that determination of success now" (ED, 2015, p. viii). In summary, it is reasonable to conclude that the expansion of common schools increased access for many, although contemporary legislative efforts to ensure school and teacher quality have demonstrated limited results.

Competency-based assessments and evaluation of programs. The requirement to pass a competency-based exam toward teaching licensure is not new, although

application of this practice is cyclical (Blackford, Olmstead, & Stegman, 2012).

Attention to the issue of teacher certification tests has waxed and waned. The launch of Sputnik in 1957 established it as a national priority, and it has remained a consistent concern from that time (2012). A review of federal legislation concerning public (K-12) and higher education compiled by the Loyola Law Review (Chopin, 2013) revealed no fewer than 16 legislative attempts aimed explicitly at increasing teacher/school quality from 1958 to 2010. Additional legislation during this span may have also included efforts toward these aims, but were not expressly stated as such. One of the provisions of NCLB required a teacher certification testing component, and while states may choose from tests such as Praxis I, Praxis II, and Praxis III (2012), there has been a nationwide increase in the use of the Teacher Performance Assessment (Goldhaber, Cowan, & Theobald, 2017). Additional examination of the Teacher Performance Assessment edTPA follows and is a focus of this study.

Statement of the Problem to be Investigated

At least 20% of new teachers entering the teaching profession will do so through an alternative route program as opposed to a traditional program (Woods, 2016). Given the increase in the development of alternative route programs to certification to address shortages in key areas, research to ensure that alternative programs prepare candidates in a manner commensurate with traditional programs as measured by edTPA is warranted.

Purposes of the Study

The first purpose of the study is to compare observed and expected pass/fail rates (categorical data) on a commercially prepared, widely adopted assessment used to ascertain candidate readiness (edTPA) between traditionally prepared and alternatively

prepared candidates. The second purpose is to determine if there is a statistically significant difference in (continuous) scores on the planning, instruction, and assessment sections between traditionally prepared and alternatively prepared candidates on the edTPA.

Formal Statement of Research Question 1

Is there a statistically significant difference between observed and expected pass/fail scores for traditionally and alternatively prepared teachers on the edTPA?

Hypothesis 1

The alternative hypothesis (H) is there is a relationship between program type (traditional and alternative route teacher preparation programs) and performance (pass/fail score). The null hypothesis (H_o) is there is no relationship between program type and pass/fail score.

Formal Statement of Research Question 2

Is there a statistically significant difference in edTPA planning, instruction, or assessment scores between participants in traditional teacher preparation programs and participants in alternative route teacher preparation programs?

Hypothesis 2

The alternative hypothesis (H) is that there are statistically significant differences in edTPA scores on planning, instruction, or assessment sections based on program type. The null hypothesis (H_o) is that there is no statistically significant difference in edTPA planning, instruction, or assessment scores based on program type.

Significance of the Study

Competency-based examinations are rooted in behaviorism, specifically the process of collecting data as to what individuals could do (what can be observed, outside of the “black box”). A behaviorist construct to learning would suggest that the external environment is the primary mediator for an individual and that this external environment is what should be noted and copied (Strauss, 1993).

Beyond the behaviorist focus on what can be seen and measured, the emphasis on high-stakes summative measurement of prescriptive skills and concepts is consistent with a knowledge-centered emphasis on curriculum (Hirsch, 1987); in this case, teacher preparation curricula. Hirsch (1987) identified the goal of the knowledge-centered curriculum as consistency in what is “known” to level the playing field for all and to prepare citizens on a national and global level. Again, while Hirsch (1987) directed his comments toward curricula for students, his emphasis on consistency, quality, and accountability are consistent with the stated goals of NCLB and Race to the Top for both students and their teachers.

As a practical matter, research comparing traditional and alternative route teacher preparation programs is lacking (Buchanan, Lang, & Morin, 2013), as will be revealed in the Review of Literature. While such programs may be compared using a variety of factors, the proliferation of the edTPA throughout the United States provides us with a common tool to compare programs that are often entirely different in their design. Comparisons of edTPA results between students enrolled in traditional and alternative route programs can be meaningful both within and between institutions provided program reviews follow when differences are indicated.

Terminology and Definitions

The ED (2004) has identified ten distinct descriptions of alternative route programs to certification. Working definitions for traditional and alternative route programs are utilized in the Review of Literature to both organize and interpret the available studies.

It is appropriate to define what is considered a traditional or alternative route program for the purposes of this study. The urban, mid-sized, private university utilized in the present study provides “traditional” teacher preparatory programs. These are four-year programs and are designed to provide undergraduates with state certification commensurate with completion of the program. Most teacher preparation coursework is on campus and occurs in the junior and senior years. Practicum experiences typically last one school year (during the final year of the program) as the candidate assumes increased responsibility for classroom instruction.

This university also offers alternative routes to certification programs. The first option is for candidates who hold a bachelor’s degree from an accredited program and wish to pursue the coursework required for state certification in one or more areas without earning an advanced degree. Coursework may be online, on campus, or a combination of both and the program requires about one year to complete. Candidates must complete the coursework, an extended practicum, and pass the edTPA. A second alternative route option is similar to the first program with an important difference. This program allows those holding an existing bachelor’s degree to pursue teacher certification commensurate with a master’s degree. These candidates must complete additional coursework, a practicum, and pass the edTPA. This program typically requires five quarters to complete.

Organization of Subsequent Chapters

This dissertation is organized to proceed from a general introduction of the topics under review toward greater specificity in following chapters. The Review of the Literature (Chapter Two) includes a review of available literature on comparisons between traditional and alternative programs utilizing a variety of aspects and concludes with an in-depth analysis of comparisons using competency-based tests (which is the focus of this study). Chapter Three includes detailed descriptions of participants, sampling procedures, instruments used, data analysis, variables, procedures, limitations, and delimitations. Chapter Four consists of an examination of the results of descriptive and other statistical analyses. Chapter Five provides a synthesis and analysis of the findings as well as recommendations for future research.

Chapter Two

Review of Literature

Is there a teacher shortage in the United States? Most laypersons would likely offer a strongly worded affirmative response to this question. In contrast to this perception, teacher production in the United States is estimated to increase 29% before the year 2022, resulting in an overall gain in the number of elementary and secondary teachers of 12% and a reduction in student-to-teacher ratios (Aragon, 2016). Moreover, despite ongoing reports of an attrition rate of 50% and the oft-cited statistic that half of teachers leave the profession within five years, federal data reveal an attrition rate of 17%, with about half of these persons planning to return to the profession following pregnancy, child-rearing, or involuntary transfers (2016). Sucher, Darling-Hammond, and Carver-Thomas (2016) reported that just under a third (between 28% and 32%) of teachers (including those who left for child-rearing or other personal reasons) come back within five years. Differences in perspective as to the scope of the problem exist; although most agree that shortages exist and will persist.

A “first blush” review of these statistics belies a complex reality faced by schools. While the above statistics may seem encouraging, further analysis of data reveals severe teacher shortfalls in rural, urban, and high-poverty areas, as well as overall shortages in areas such as special education, science, and mathematics (Aragon, 2016; Hussar & Bailey, 2016). In 2015-16, 48 states reported that special education was their greatest area of shortage (Sucher, Darling-Hammond, & Carver-Thomas, 2016). Mathematics and science teachers followed close behind: 42 states reported shortages in these areas (2016).

In response to this deficit, school districts, universities, and other institutions have developed programs providing alternative routes to teacher certification. Although initial programs were designed by school districts, current programs are more often sponsored by institutions, with school districts as partners (Basinger, 2000). Current estimates suggest that 20% of new teachers entering the profession do so through an alternative certification program (Woods, 2016). Alternative route programs have secured support from the ED, and are hailed as an innovation (ED, 2004). Given the widespread support for these programs and persistent need for teachers in key areas, these programs are likely to persist as ongoing efforts to address shortages. As such, this study represents an attempt to evaluate these “alternative” programs as to their efficacy when compared with traditional teacher preparation programs.

Definitions

Alternative route programs are diverse in their design because of differences in regulations, available funding, and the interests of stakeholders that develop the programs (Mitchell & Romero, 2010). As a result, the term *alternative routes* describes a vast array of program models encompassing everything from informal support provided to teachers with emergency teaching permits to well-organized programs that culminate with full certification and an advanced degree (ED, 2004). Various professional organizations have endeavored to classify alternative programs resulting in up to 10 distinct descriptions (2004). An examination of specific programmatic differences among alternative route programs is beyond the scope of this work. For the purposes of this literature review, a working definition provided by Mitchell and Romero (2010) will suffice: a traditional program is one requiring coursework and a student teaching

experience prior to becoming certified and obtaining a bachelor's degree. In contrast, an alternative program includes a residency, coursework (online, in person, or hybrid), and may offer financial incentives toward certification with or without the opportunity to earn an advanced degree. Most alternative programs require applicants to possess an existing bachelor's degree upon application to the program (2010). Chapter Three will follow with specific program descriptions and sample characteristics for the purposes of this study.

If a definition of “alternative” is elusive, an understanding of what constitutes “effectiveness” when comparing traditional and alternative teacher education programs is even more complex. Claims of efficacy for alternative and traditional preparatory programs have centered on aspects of pedagogy, self-efficacy, and student achievement (Tournaki, Lyublinskaya, & Carolan, 2009). Candidate satisfaction, retention (Zhang & Zeller, 2016), and aspects of multicultural awareness (Brown, 2005) are also cited as evidence of effectiveness. This literature review is organized through an examination of these themes, and critical analysis is embedded throughout.

Measures of Effectiveness

Increased multicultural awareness. Research (qualitative or quantitative) attempting to compare differences in the multicultural awareness of traditional or alternative route candidates is lacking; challenges associated with common definitions and adequate measures of multicultural awareness may be a contributing cause. Nevertheless, two qualitative studies conducted on alternative route programs claim increased cultural awareness as a result of these programs; bold claims from studies with qualitative designs. A qualitative study conducted with 10 total native and non-native

English speaking para-educators participating in an undergraduate teacher licensure program sought to identify pedagogical approaches and perceived benefits of participating in their teacher preparation program (Burbank, Bates, & Schrum, 2009). Four themes emerged through a continuous comparative analysis of survey data, interviews, and course assignments (2009). The first two themes related to seminar content and pedagogical approaches. The third and fourth themes were the unique perspectives of immigrant paraprofessionals and the benefits of participation for native English speakers. These themes were cited as evidence of increased multicultural awareness (2009). The authors highlighted participant contributions that indicated feelings of validation for their unique perspectives and contributions suggesting that they had gained “insider information” from fellow students in the program, even if they did not share these experiences (2009). It is clear that participants in this program believed they developed increased cultural understanding, although it is impossible to extend such claims to other alternative programs. While not commensurate with the stated purpose of the study, the authors reported that 4 of 10 participants attained certification of some kind (2009). It is possible that their program may hold potential toward increased workforce diversity given that a portion of participants were non-native English speakers, although the number of non-native English speakers that attained certification is unknown.

A second qualitative study sought to investigate the role of service learning in an alternative route program with four goals: increased cross-communication skills, inclusive behaviors toward diverse groups, insight into educational policies, and ability to generalize these skills to classroom practices (Brown, 2005). The study included 73 graduate teacher candidates enrolled in an alternative route program assigned to a

diverse, urban school over a five-year period (2005). Participants were required to propose, negotiate, and complete a service learning project that benefitted the school, reinforced their content knowledge, and met five characteristics of service learning (2005). Reflective journals, reports summarizing each individual's service learning project, and discussions were the instruments used to determine the projects' effectiveness toward meeting the four program goals (2005). A specific methodology used to code and analyze data was not disclosed, although commentary from individuals was included and discussed at length. The author claimed, "The results of this five-year research study indicate that service-learning embedded in the restrictive time frame of a one-year graduate teacher education program can positively influence the multicultural perceptions, cross-cultural communication skills, and social justice cognizance of future teachers" (p. 69). While the questions asked in this study are significant, it is unfortunate that the author did not conduct an analysis of the volume of artifacts collected or disclose her methodology, as claims made through carefully designed qualitative studies can further inform the profession and raise important questions (Gall, Gall, & Borg, 2007). The author's failure to report these aspects of her study further weaken her overly general claim regarding the impact of service learning on cross-cultural awareness and social justice cognizance.

Candidate satisfaction. Candidate satisfaction with his or her preparatory program can be reasonably understood to include the concept of self-efficacy, which is the focus of a great deal of current educational research. While these studies abound, studies comparing graduates from alternative route programs and traditional preparatory programs are scarce. The review of the literature revealed three such studies. The first

study included six research questions focused on identifying interaction effects between the type of program (alternative or traditional), years of teaching experience, and perceived self-efficacy toward six job-related domains as measured by a survey (Lowe, 2012). The six job-related domains identified were: planning and preparation for learning, delivery of instruction, assessment and follow up, classroom management, family and community outreach, and professional responsibilities (2012). One hundred and four participants had five or fewer years of teaching experience and were employed by districts in Louisiana or Mississippi (2012). Statistical analysis of survey results utilizing a two-way ANOVA revealed no statistically significant differences between alternative and traditional programs within any of the six job-related domains identified in the study (2012).

The second study employed a purposeful sample of 288 first year, newly employed teachers from 50 school districts in the greater Houston area (Fox & Peters, 2013). The researchers utilized a two-tailed independent *t*-test to ascertain if statistically significant differences in teacher efficacy scores as measured by the Teachers' Sense of Efficacy Scale existed between graduates of traditional versus alternative route programs (2013). Analysis of the scores revealed no statistically significant differences between groups, although it is interesting to note that the sample utilized in the study included new teachers defined as "all those who were newly employed with the district and did not have any teaching experience" (Fox & Peters, 2013, p. 6). While this study utilized similar statistical procedures as the first, it is important to recognize that results were similar between respondents with 0-5 years of teaching experience and those who had no experience. While one might expect that teachers who have yet to assume responsibility

for their own classrooms might have an inflated sense of self-efficacy regardless of their certification pathway, this does not appear to be the case.

Buchanan, Lang, and Morin (2013) conducted a mixed methods study which addressed perceived differences between traditionally prepared and alternatively prepared teachers serving in grades 9-12 in Tennessee. Quantitative analysis (causal-comparative and correlational) was conducted through the use of a survey sent to 141 teachers and returned by 34. The survey was based on characteristics included in Tennessee's teacher evaluation rubric and included items adapted from other surveys. Further analysis was conducted through interviews with a selected number of those that returned the voluntary survey. A number of hypotheses were included in the study that while interesting, were beyond the focus of this literature review, including principal perceptions and hours of coursework (2013). The investigator's analysis of this study focused on teacher satisfaction or self-efficacy.

The authors elected to limit their sample to secondary teachers in grades 9-12. They asserted that secondary teachers included the highest number of alternatively prepared teachers in Tennessee (Buchanan et al., 2013). This purposeful survey limits generalization, as does the response rate of 24%, although Fowler (2009) acknowledged that there is no universally accepted minimum response rate for a survey. Moreover, bias may exist if survey respondents have a particular interest in the subject of the research (2009). Despite limits toward generalizability of findings and potential bias by survey respondents, data revealed no statistically significant relationship between the amount of teacher training, credit hours, type of preparatory program (traditional or alternative), and self-efficacy as measured by their survey (2013).

Candidate retention. Zhang and Zeller (2016) conducted a mixed methods study designed to identify differences in short and long-term retention of teachers from three different programs in North Carolina. The authors defined the first program type as “regular,” that is, undergraduate preparation programs that concluded with a bachelor's degree and state certification. The second was called a “lateral entry alternative program” and was described by the authors as a sink or swim program in which candidates holding a bachelor's degree began teaching immediately and were expected to complete courses required for teacher certification within a specified period of time. The third program (also an alternative route program) was called the NC TEACH program and was based in North Carolina. This program targeted mid-career professionals and was designed to “ease” these candidates into the profession. Five weeks of intensive summer instruction was required before the onset of the subsequent school year. Some candidates who entered the program had already been employed as teachers for one year; others were about to begin serving in teaching positions. All participants in this program were expected to complete 12 semester hours of graduate work in the school year immediately following their intensive summer session (2016).

Zhang and Zeller's (2016) study utilized a survey comprised of 22 questions adapted from an unpublished tool used in a similar study by Johnson and Birkeland (2003). A total of 60 educators (20 from each of the three programs described) completed the survey at year two, year three, and year seven of their employment. Qualitative analysis included coding of responses by two researchers, and interrater reliability was reported at .89. Proportionately more lateral entry teachers reported being less prepared to teach in the way that they felt they were expected to teach. Age, gender,

ethnicity, school level, marital status, and parents' occupation did not appear to influence retention (2016).

Quantitative analysis consisted of a logistic regression model used to predict the outcome variable of retention by the explanatory variable of program type (Zhang & Zeller, 2016). Findings revealed that statistically significant differences exist in the short-term (years two and three) and the long-term (year seven). Teachers in the lateral entry alternative program had lower retention rates than traditionally prepared teachers or those in the NC TEACH alternative program in both the short and long-term (2016). The authors reported that of all variables examined in the study (years teaching, having children, ethnicity, gender, teaching assignment, marital status, parents' occupation, and type of preparation program), the type of preparation program was the only variable with predictive validity toward retention in the field. Zhang and Zeller (2016) qualified their findings in the following way:

Teacher retention likelihood partially depends on the type of preparation teachers receive. Although the type of teacher preparation is an important factor that explains teacher retention, predicting retention is more complex and will require further consideration and examination of other factors. (p. 86)

Analysis of these findings may be misleading without careful examination of the program descriptions provided by the authors of the study. Initially, there seems to be a real distinction between the lateral program in which teachers are employed by districts and expected to complete coursework within a specified period of time, and the NC TEACH program in which participants completed a five-week summer session, followed by employment in public schools. Certification was commensurate with completion of

coursework. These programs shared similar characteristics; both required coursework while teaching and both required that candidates have a bachelor's degree prior to beginning the program. While retention of teachers from traditional programs remained consistent over the long term, retention of teachers from *both* the “lateral” and NC TEACH programs declined over time (Zhang & Zeller, 2016). Nevertheless, this study provided evidence that differences in retention between traditional and alternative programs utilized in North Carolina exist and further research is warranted.

Grissom (2008) conducted a study utilizing a national sample of 4,000 public school teachers. The purpose of the study was to determine if teachers who enter the profession through alternative route to certification programs leave the profession at higher rates than do teachers from traditional programs. The study utilized the 2003-2004 Schools and Staffing Survey (SASS) and the 2004-2005 Teacher Follow-up Survey (TFS). The SASS is administered to a nationally representative sample of teachers every four years to identify teacher shortages, teacher characteristics, and school characteristics. The SASS asks that teachers identify the type of program they participated in or their employment status. The choices from which the teachers may select are

- regular or standard,
- probationary,
- provisional,
- temporary,
- emergency waiver, or
- other (2008).

The author clarified that teachers identified as probationary had met all certification requirements except for the completion of the probationary period (Grissom, 2008). These teachers, along with those that classified themselves as having completed a regular or standard program, were considered traditionally prepared teachers (2008).

The TFS is administered the year after every SASS administration to a subset of those that participated in the SASS (Grissom, 2008). One of the purposes of the TFS is to identify mobility in the teacher workforce, and it is used to analyze turnover between the two-year period from SASS to TFS administration (2008).

Multivariate analysis of program type, retention, and school characteristics revealed that teachers from alternative route programs were less likely to remain in their current positions (82.3%) after one year than traditionally prepared teachers (85.6%) (Grissom, 2008). The author stipulated that although the difference is statistically significant ($t = -1.82$), it was not especially large. It is important to note that the difference described above is a difference in “moving,” not in remaining in the profession. Teachers from alternatively prepared programs were less likely to stay in their initial positions, but the difference between traditional and alternative program teachers in leaving the profession was less than one-half of one percent and was statistically insignificant. The author concluded that although teachers from alternative route programs were less likely to stay in their initial position from year to year, this may have been due to characteristics of the schools they work in (e.g., urban school, rural school, elementary school, secondary school, public charter school, etc.) rather than the attributes of the program the teachers participated in to earn certification (2008).

Grissom's (2008) findings suggest that previous research comparing attrition between traditionally prepared and alternatively prepared teachers may be overly simplistic if the research designs utilized cannot differentiate between those that move to another position or leave the profession altogether (2008). Moreover, as many alternative route programs are designed to address shortages in urban schools, examining overall retention between program types without considering school characteristics does little to inform the field if alternative programs are addressing one of their stated objectives (2008).

Grissom's (2008) findings appear to contradict those by Zhang and Zeller (2016). The sample size, sample characteristics, and statistical analysis used in Grissom's (2008) study tended toward greater generalization of findings regarding retention than does the study by Zhang and Zeller (2016), which was limited by sample size and sample characteristics. Unfortunately (as acknowledged by the author), Grissom's (2008) study did not include questions designed to help identify differences in the types or characteristics of alternative route teacher education programs.

Student achievement. Given the current emphasis on accountability in public schools, it is not surprising that studies examining the relationship between teacher certification pathway and student achievement are a focus. Boyd, Grossman, Lankford, Loeb, and Wyckoff (2006) conducted a quantitative study in the state of New York using regression analysis designed to ascertain the relationship between student achievement (as measured by state tests aligned with state learning standards) for students in Grades 3 through 8 (mathematics and English language arts) and the certification pathway. The model included fixed effects for years, grades, and schools (Boyd, Grossman, Lankford,

Loeb, & Wyckoff, 2006). The authors reported that effects of teacher pathways on achievement differed by grade. They offered a model that grouped students into two groups: one comprised of fourth and fifth graders and another consisting of sixth, seventh, and eighth graders (2006). Results indicated that students taught by Teacher Fellows (an alternative route program) made significantly greater gains in elementary mathematics achievement than did students taught by teachers from traditional programs, and that there was no statistically significant difference between scores for students taught by Teach for America teachers and students taught by traditionally prepared teachers in mathematics (2006). Elementary student achievement in English Language Arts for teachers prepared by alternative route programs did differ from student performance by students with traditionally prepared teachers, although these differences disappeared by the third year (2006). Students of middle-level alternatively prepared teachers performed at least as well as or better than students of traditionally prepared teachers in measures of mathematics achievement, and there was no statistically significant difference in students' English language arts performance between alternatively and traditionally prepared teachers. In summary, data indicated that the academic performance of students taught by alternatively prepared teachers in Grades 3 through 8 may differ initially (with differences resolving over time) or is commensurate with the performance of students with traditionally prepared teachers. Students of alternatively prepared teachers were estimated to show gains .02 lower than traditionally prepared teachers in mathematics and .03 less in English Language Arts. This gap in student performance narrows in an alternatively prepared teacher's third year; students in this group improved 5.6 percent of a standard deviation more than the performance of

students with traditionally prepared teachers. The rigorous process used for matching students with teachers and teachers with programs along with the magnitude of student data collected increases the investigator's confidence in these results, as does the thoroughness of the statistical analysis provided – e.g., significance was tested and reported at three levels, and sample sizes exceeded the minimum recommendation of 10 per predictor (Field, 2013).

A similar study conducted in Florida sought to explore the relationship between teacher certification program (traditional or alternative route) and student achievement (Sass, 2013). The study utilized a value-added regression model controlling for school effects (Sass, 2013). The first alternative route program was district centered and provided online instruction and mentoring while candidates served as teachers. This program was called the Education Preparation Institute. Candidates in this program earned certification through one year of non-transferrable coursework (online and face-to-face), completion of a practicum experience, and passing scores on standard certification exams. Participants in the other alternative route programs were required to pass competency tests (either exams administered by the American Board for Certification of Teacher Excellence or a subject area certification exam) with no additional coursework. The third group consisted of traditionally prepared teachers (2013).

Sass (2013) reported that student test scores were normed by grade and year so coefficient estimates could be interpreted as standard deviation units of student achievement. Reporting results in terms of student achievement is consistent with the context he used to frame his findings. Overall results cited by the author indicated value-

added estimates for district alternative programs was 1 to 2% of a standard deviation greater than estimates for traditional programs (Sass, 2013). In contrast, the scores of Education Preparation Institute teachers were 2 to 4% lower than traditionally prepared teachers (2013). Performance between teachers who were only required to pass a competency test was significantly different from those from traditional programs in mathematics; the test-taking group outperformed traditionally trained teachers by 6 to 8% of a standard deviation (2013). While these findings lend support to the efficacy of alternative programs in relation to student achievement, it is important to note that this study is a working paper published by the Andrew Young School of Policy Studies and findings have not been vetted through a peer-reviewed journal; tables referenced in the text were not included for examination.

Clark et al. (2013) conducted a study using mixed methods to compare middle and high school student mathematics achievement under alternative route teachers (Teach for America, Teaching Fellows) versus teachers from traditional programs. The study included teacher participants from a sample including 11 states, 15 districts, 82 schools, 287 classrooms, and 287 teachers (2013). Students who attended the same school and were enrolled in the same mathematics course were randomly assigned to teachers who were alternatively trained or to comparison classrooms with traditionally trained teachers (2013). The authors reported that random assignment methods used in their study lent credence to their results. Utilization of a regression model revealed that students of teachers from the Teach for America group had statistically significantly higher mathematics achievement scores when compared with traditionally trained teachers while no differences in achievement were identified between Teaching Fellows teachers and

those from traditional programs (2013). Nonexperimental analysis including analysis of educational background, performance on mathematics knowledge tests, and professional development experiences revealed that factors other than the route to certification may have been associated with student achievement (2013). This study provides additional evidence regarding the efficacy of alternative programs beyond previously discussed studies utilizing a specific state, as 11 states from across the nation were represented.

A study conducted by Glazerman, Mayer, and Decker (2006) utilized a smaller sample representative of six urban regions. The study yielded similar findings relative to mathematics achievement and Teach for America (TFA) teachers, despite no differences in reading achievement. Students in the same schools and at the same grades were assigned to Teach for America teachers or to non-TFA teachers (control teachers). Students were randomly assigned to teachers to avoid any intentional or unintentional bias in how students were assigned. The impact of TFA teachers on student mathematics achievement ranged from .13 to .19 standard deviation units and was always statistically significant; the comparative impact of TFA teachers on student reading achievement was not statistically significant when compared with non-TFA teachers when a regression model was utilized (Glazerman, Mayer, & Decker, 2006).

A significant limitation of this study is the lack of specificity in how the control group (non-TFA teachers) was defined. The control group included traditionally certified, alternatively certified, or uncertified teachers; any teacher who had not participated in TFA at any time (Glazerman et al., 2006). In light of this point, this study does not provide strong evidence favoring TFA as a robust alternative program. This is especially so since the comparison group included teachers with divergent backgrounds;

including teachers from other alternative route programs. Glazerman et al.'s (2006) research question, "do TFA teachers improve (or at least, not harm) student outcomes" (p. 77), suggests that they understood the limitations of their findings and that at best, despite statistically significant gains in mathematics achievement, the TFA program can only claim to have done no harm.

Performance on competency-based assessments. Tournaki, Lyublinskaya, and Carolan (2009) conducted a study designed to identify the relationships between professional pathway (alternative or traditional route to certification) and effectiveness, as measured by the Danielson Observation Scale, as well as efficacy as measured with the Teacher Efficacy Scale. The sample included 66-71 graduate students in their last semester of coursework attending a public university in New York City (2009). A total of 26 students participated in traditional teacher preparation programs, and 57 students participated in alternative programs. A one-way analysis of variance revealed no statistically significant relationship in either effectiveness or efficacy based upon the instruments administered. These findings must be interpreted with caution. The authors acknowledged that the teachers in the sample had less than five years of teaching experience, and therefore, may have had a "fixed" sense of their self-efficacy and effectiveness that was not dependent upon their teacher preparation program. The addition of a control for initial efficacy/effectiveness perceptions would have enhanced the design of the study (2009).

Reflection on the practical and statistical significance of Goldhaber, Cowan, and Theobald's (2017) work requires a contextual understanding of the development, characteristics, and overall utilization of the competency-based assessment utilized in

their study, the Educative Teacher Performance Assessment (edTPA). Researchers at Stanford University developed the edTPA with subsequent revisions provided by American Colleges of Teacher Education (AACTE) and Evaluation Systems (Goldhaber et al., 2017). Evaluation Systems is a member organization of Pearson Evaluation Group, and Pearson oversees the administration and scoring of the test at this time (2017).

The edTPA is a subject-specific performance assessment that evaluates a common set of teaching principles and instructional strategies that are focused on specific content learning outcomes for preschool through twelfth-grade public school students (Pecheone, Whittaker, & Klesch, 2017). There are 27 versions of the edTPA specific to different content specializations (Goldhaber et al., 2017). Each version includes 15 different rubrics, each of which is scored from 1-5 and weighted equally. The overall sum of scores may range from 15-75 (2017).

Washington State groups the rubrics into three categories: planning, instruction, and assessment (Goldhaber et al., 2017). Three additional rubrics (student voice) are required in Washington State but are not included in the candidate's score at this time (2017). Professional Educator Standards Board's (PESB) edTPA student voice committee recommendations (2016) included references to ongoing data collection on the student voice rubrics and mentioned the possibility of this section becoming "consequential" to candidates in 2018. The committee did not provide detail as to underlying discussions that informed their recommendations (2016).

Completion of the edTPA assessment is viewed by some as similar to the process for National Board Certification. While National Board Certification is esteemed as recognition of teaching excellence to those in the field, it is possible that many public

school administrators are less familiar with the requirement to pass the edTPA or the scope of work required to do so. Participants are required to prepare an instructional plan ranging from three to five hours of active instruction (Kroeger & De Vares, 2014). All instructional materials must be created or acquired in accordance with the plan. Copies of student assignments and assessments are collected along with feedback on student work from selected students. The delivery of instruction is submitted as unedited video on a secured website as are all other aforementioned materials (2014).

edTPA scores are reported as “pass/fail” based upon a cut score (Pecheone et al., 2017). Pearson provides recommendations as to cut scores, but individual states have autonomy to adjust scores as they see fit. Washington State is no exception. No fewer than four revisions to the criteria required to pass have occurred from September of 2011 when the State of Washington Professional Educator Standards Board determined that portions of the edTPA would be consequential to candidate certification (Professional Educator Standards Board, 2018). Washington State’s criteria to pass in each of the 23 content specializations have increased through the revisions, although the current score required to pass in most areas (40) is less than the overall national mean in 22 of 23 teaching fields (Pecheone et al., 2017).

Pecheone, Whittaker, and Klesch (2017) reported that 13 states have a policy in place that requires a state-approved performance assessment in order to earn state licensure or program completion. Careful examination of the explanatory data revealed that the edTPA was “one of many” assessments approved for this purpose (2017). While it is the reader’s responsibility to “read the fine print” when information is presented visually (in this case, a map of the United States) depicting usage of the edTPA, this

portrayal of data is misleading in that it leads the reader to assume that edTPA is the only approved assessment in these states.

Three states (Alabama, Ohio, and Connecticut) are reported to be taking steps toward implementation of a required test (Pecheone et al., 2017). Careful examination of the key provided includes a disclosure that edTPA is, once again, one of the assessments being considered (2017). No information is provided as to each state's process; some may be conducting an initial investigation into the usage of exams for licensure while others may be nearing the adoption of a policy requiring candidates to pass a test (2017).

Twenty states are reported to be participating in the edTPA assessment (Pecheone et al., 2017). Again, examination of this claim revealed that "at least one institution" in the state is exploring or piloting the edTPA (2017). Again, the portrayal of these states as "participating in the edTPA" given the caveats provided is misleading. Despite the investigator's misgivings regarding the portrayal of edTPA implementation/adoption data, it is widely utilized in many forms. It is required in the State of Washington, and the investigator hopes this overview of the edTPA assists in the interpretation of the study by Goldhaber et al. (2017) that follows.

Goldhaber et al. (2017) recently published a study of particular note. Although they did not undertake an analysis of edTPA performance comparing traditional and alternatively prepared teachers, they did utilize longitudinal edTPA data to provide estimates of the predictive value of the test toward employment in Washington State public schools and teaching effectiveness (2017). The question of predictive value based upon pass/fail analysis of data was not surprising; a passing score is highly predictive of employment in a Washington State public school the following school year. This reflects

Washington State's requirement that candidates pass the exam prior to certification in our state. Pass/fail data were also related to student reading performance (i.e., Measures of Student Progress scores for grades three through eight), but not with mathematics performance. Continuous edTPA scores were related to teaching effectiveness in mathematics (Measures of Student Progress scores for grades three through eight), but not with reading (2017).

The authors' analysis of pass/fail rates among minority candidates was most disturbing (Goldhaber et al., 2017). Hispanic candidates in Washington State were more than three times more likely to fail the edTPA when compared with non-Hispanic White candidates. While these candidates may have been unlikely to secure teaching positions regardless of the requirement to pass a high-stakes test due to other (unknown personal characteristics), this may negatively influence Washington State's efforts to further diversify the workforce (2017). Again, while the study did not address comparisons between the type of teacher education program that candidates attended, it reinforces the need for further analysis of the edTPA as used in our state toward candidate licensure and evaluation of teacher education programs.

Concerns regarding the edTPA. The investigator would be remiss if she failed to acknowledge concerns in the field regarding the authorship, lack of empirical research supporting the tool, and potentially misleading claims proffered by the current administrators of the edTPA. Hebert (2017) began her critique of the edTPA by citing evidence of a conflict of interest in that much of the evidence directly related to the edTPA was written by the authors of the edTPA themselves. A plethora of additional "supporting research" is also included on Pearson's website. Closer examination

revealed an excess of 200 articles which were organized by the subtopics of the edTPA but did not include any explanation as to how the 200 articles related to the edTPA (2017).

Hebert (2017) also expressed concerns regarding the use of the Performance Assessment for California Teachers (PACT) as evidence in support of the edTPA. PACT is often referred to as the precursor to the edTPA, yet important distinctions between the tools exist such that shared claims of efficacy are inappropriate (2017). PACT is aligned to a specific set of state standards, which suggests that PACT may be valid when used in this context. The edTPA has not provided evidence that the test aligns with other state standards. Moreover, there are no national standards with which edTPA could align. The Model Core Teaching Standards developed by the Interstate Teacher Assessment and Support Consortium (InTASC) are referenced by AACTE and the Stanford University Center for Assessment, Learning, and Equity (SCALE), yet InTASC publishes links to a variety of resources specific to individual states and their standards (2017).

A final point of contention provided by Hebert (2017) are the practical differences in how PACT and edTPA disseminate information about the tests and how the tests themselves are scored. While scorers for PACT and edTPA must be similarly qualified, PACT scorers reside locally, and the tests are scored locally. This is of practical significance because those who are scoring are likely familiar with state initiatives, local teacher education programs, and curricula. The edTPA is scored nationally with the assertion that this practice “breaks down barriers” and assures that only the highest qualified individuals are selected to score. Interestingly enough, Pearson is investigating the provision of a regional scoring pool which would allow teacher education programs

some voice in how/who scores these assessments (2017). Scorers of the edTPA and other stakeholders have also expressed concerns that Pearson does not provide sufficient material to support test takers and those that instruct them. Those seeking this information must rely upon the limited resources provided by Pearson. Those who score (or who have served as scorers in the past) are required to sign nondisclosure agreements prohibiting them from discussing the content of the test with others (2017).

Parkes and Powell (2015) extended many of Hebert's (2017) claims and provided additional concerns. One further concern related to the lack of feedback that candidates receive after taking the edTPA. Participants receive a numerical score indicating that they have passed or failed, but do not receive any additional written feedback, a rationale for the scoring, or suggestions for improvement (2015). In light of Hattie's (2012) meta-analysis establishing explicit feedback as an evidence-based practice, this seems counterproductive; particularly when retakes (a single task, multiple tasks, or the entire assessment) of the edTPA are allowed (Pecheone et al., 2017).

The considerable expense incurred by pre-service teachers is a second concern raised by Parkes and Powell (2015). Pearson is a for-profit company, and the test is both administered and scored by them. The test currently costs \$300.00 per administration. Candidates who need to retake sections are subject to additional charges between \$100.00 and \$300.00. Scorers of the edTPA are paid \$75.00 for each scoring opportunity which seems disproportionately low given the cost to candidates. In response to this expense, several higher education institutions are charging a "lab fee" spread over several courses and applied toward a budget that is then used to pay for a candidate's first attempt. While

this may assist in spreading the financial burden over time, the candidate still bears the final cost (2015).

Summary

Despite projections that the overall number of teachers will increase by 29% by the year 2022 (Aragon, 2016; Hussar & Bailey, 2016), shortages will continue to exist in critical areas such as special education, science, and mathematics (2016). In attempts to ameliorate these shortages, policy makers, teacher preparation programs, and school districts have designed and implemented a diverse range of alternative route to certification programs (Basinger, 2000). The ED (2004) has identified at least 10 distinct alternative program models implemented by stakeholders as they grapple with ongoing shortages.

As reliance on alternative route programs increases, research attempting to compare their effectiveness with traditional programs is in its infancy, and few studies exist (Buchanan et al., 2013). Nevertheless, alternative route and traditional programs may be compared by factors such as multicultural awareness, retention, candidate satisfaction or self-efficacy, student achievement, and performance on competency-based assessments.

Appreciable differences in multicultural awareness in candidates who participate in traditional or alternative route programs cannot be ascertained at this time due to a lack of much-needed research, including an apparent absence of quantitative and mixed methods studies. Two qualitative studies reviewed cited overall increases in multicultural awareness but were focused on specific programs at single sites and the usual caveats

regarding the limits of descriptive findings must be observed (Brown, 2005; Burbank et al., 2009).

Research addressing the construct of self-efficacy abounds. Unfortunately, there are a limited number of studies comparing candidate satisfaction or self-efficacy between traditional and alternative route programs in the body of research. Review of research findings were consistent in that there were no statistically significant differences between teachers trained in alternative or traditional programs (Buchanan et al., 2013; Fox & Peters, 2013; Lowe, 2012).

With regard to teacher retention between traditionally prepared and alternatively prepared teachers, it was clear that differences between long-term longevity in the profession exist between traditionally prepared and alternatively prepared teachers (Zhang & Zeller, 2016). Traditionally prepared teachers demonstrated consistent, high rates of retention in the short and long term. Having reported this, the authors also acknowledged the disparity of research between traditional teacher education programs with regard to retention and other factors and the limited amount of research available regarding alternative route programs (2016). Findings by Grissom (2008) revealed that teachers from alternative route programs were less likely to remain in their current positions (82.3%) after one year than traditionally prepared teachers (85.6%). The author stipulated that although the difference is statistically significant ($t = -1.82$), it was not especially large and emphasized that the difference was a difference in “moving,” not in remaining in the profession. Teachers from alternatively prepared programs were less likely to stay in their initial positions, but the difference between traditional and

alternative program teachers in leaving the profession was less than one-half of one percent and was statistically and practically insignificant.

Comparisons between alternative and traditionally prepared teachers with respect to student achievement were mixed. Boyd et al. (2006) compared alternative route and traditional programs and reported that students taught by teachers from an alternative program made greater gains in elementary mathematics achievement than those from traditional programs. Alternatively prepared teachers also demonstrated greater student gains in English language arts achievement, although these gains disappeared by the third year (2006). Clark et al. (2013) conducted a similar study and concluded that students with teachers from Teach for America (an alternative route program) had statistically significantly higher mathematics achievement scores when compared with students of teachers from traditional programs. The authors were conservative in their discussion and recommended further research to ascertain if these differences were due to participation in professional development or performance on mathematics knowledge tests (2013).

Research comparing candidate performance on competency-based assessments is sorely lacking. One study compared differences in performance between alternatively and traditionally prepared teachers on the Danielson observation scale (Tournaki et al., 2009). Analysis revealed no statistically significant differences in performance between groups (2009).

The study by Goldhaber et al. (2017) is of particular importance although it was not a comparison between program types. This study was conducted in Washington State using longitudinal MSP and edTPA data. Successful performance on the edTPA was

determined to be highly predictive of employment in a public school. This is not surprising given that a passing score is required for licensure and subsequent employment in a public school in Washington State. There was a relationship between categorical pass/fail data and student reading achievement, but not with mathematics achievement. When continuous edTPA scores were analyzed, they were related to teaching effectiveness in mathematics, but not in reading.

This review of key findings in the literature has further established the lack of research in teacher education in general (Knight et al., 2012), expressly, research comparing traditional versus alternative route programs (Buchanan et al., 2013). The paucity of research comparing traditional versus alternative route programs with respect to performance on the edTPA is particularly surprising given the increased usage of the measure despite ongoing controversy among experts in the field (Hebert, 2017; Parks & Powell, 2015).

Higher education teacher preparation programs do not have access to edTPA data other than a report listing an individual candidate's scores (Parkes & Powell, 2015). Pearson does not disaggregate results by program type (alternative route or traditional program) or chooses not to publish the information if they do. The State of Washington Professional Educator Standards Board provides summaries of edTPA results by program and teaching content area. Unfortunately, a program is interpreted as a university or college and not a type of certification program (traditional or alternative).

Unfortunately, individual teacher preparation programs will need to undertake this analysis if it is to occur. Research comparing results between alternatively and

traditionally prepared teachers on a nationally recognized competency test (edTPA) is both timely and addresses a gap in the literature.

Chapter Three

Methodology

Methods of Inquiry and Rationale

This study utilized a deductive approach consistent with a postpositivist paradigm. The postpositivist approach assumes an objective reality, albeit one that can only be imperfectly measured (Gall et al., 2007). This approach also assumes that reality can be measured and that associated variables can be identified and measured in relation to each other (2007).

A deductive approach requires that the researcher (investigator) assume etic methods, methods which require that the researcher maintain the role of an outsider in relation to the phenomena studied (Vogt & Johnson, 2011). Reality can be understood by utilizing samples to undertake observations from which theories are deduced (Rovai, Baker, & Ponton, 2014). Hypotheses are generated and tested through observation and data analysis. Conclusions flow from analysis of data (2014).

Postpositive, deductive research methods are considered quantitative research methods. Quantitative methods utilize numerical data derived from samples with the aim of generalizing to a larger population. Although the investigator has articulated the purpose of the study in previous chapters, it is appropriate to provide a brief rationale for the selection of quantitative methods of inquiry at this time. The reader is familiar with the purpose of this study: to identify whether statistically significant differences exist in edTPA performance between traditionally and alternatively prepared teachers. One of the strengths (and limitations) of all research is that conclusions are characterized by the concept of “refutability of knowledge”: a phenomenon may be considered supported by

evidence, but that it may not be considered “true” as subsequent study may provide contrary evidence (Gall et al., 2007). The investigator selected a quantitative methodology as an initial step in understanding a potential difference between program types using a relatively new assessment tool. In this sense, this study might be considered an exploratory study from which further quantitative and qualitative research may follow. Review of the information contained throughout this chapter will further substantiate the selection of quantitative procedures and associated statistical procedures.

Methodological Approaches

The research questions (and following hypotheses) were designed to identify whether statistically significant differences exist in edTPA scores between candidates prepared by a traditional or alternative route certification program. The study utilized a causal comparative design as it included an analysis of data that was already collected for purposes other than this study. Participants were observed (collection of their edTPA scores) with no alteration in their situation.

Given that edTPA data were collected as an existing requirement for program completion at the university and used by faculty for ongoing program evaluation, review by the Institutional Review Board (IRB) was not required. The university’s IRB guidelines stated, “the IRB is responsible to review and approve any proposed research with human participants that occurs outside of the established or commonly accepted educational settings involving normal educational practices such as regular course evaluations or student assessment” (Seattle Pacific University, 2017, p. 1). The investigator reviewed the research proposal and the exemption from the IRB process with the university’s IRB representative twice to satisfy due diligence.

Statistical Analysis

The selection of specific statistical tests extends from the research question, the hypotheses included in the study, and the characteristics of the data. Scores on the edTPA are reported to candidates as “pass/fail” and are based upon an averaging of subtests that are scored numerically. Research on the edTPA is limited, although Goldhaber et al. (2017) provided a thorough examination of edTPA results utilizing both non-parametric and parametric procedures. This study utilized both non-parametric and parametric procedures as well.

Non-parametric statistical analysis. Categorical edTPA data (pass/fail) were analyzed using the chi-square test of independence. The data met the following statistical assumptions for this test as recommended by McHugh (2013):

- data consisted of frequencies or counts of cases,
- there were two variables measured as categories,
- the categories were mutually exclusive,
- each subject contributed data to only one cell in the contingency table, and
- the sample size exceeded the number of cells multiplied by five.

Parametric statistical analysis. Numerical edTPA data were analyzed using Multivariate Analysis of Variance (MANOVA), which is an extension of ANOVA when multiple dependent variables exist. Data were evaluated in light of the following statistical assumptions as described by Rovai, Baker, and Ponton (2014):

- random selection of samples to allow for generalization,
- continual variables (interval/ratio) scale,
- one or more categorical variable with multiple categories,

- measurement without error,
- multivariate normality,
- no outliers present, and
- independence of observations.

Assumptions were reviewed, and although random selection of samples was not exercised, issues of normality were assumed to have been addressed through the presence of a large sample size, and other indicators which are further explored in Chapter Four.

Participants

This study included 565 undergraduate and graduate students enrolled in teacher preparation programs at a mid-sized, urban, private university in Washington State. Demographic data from the 2016-17 school year were analyzed and represents the most current (and complete) school year (Seattle Pacific University, 2018). Of 135 participants in the school of education, 119 self-identified as female and 16 as male (Seattle Pacific University, 2018). A total of eight students identified as Hispanic and the remaining (127) identified as non-Hispanic. Figure 1 depicts further analysis by ethnicity (2018).

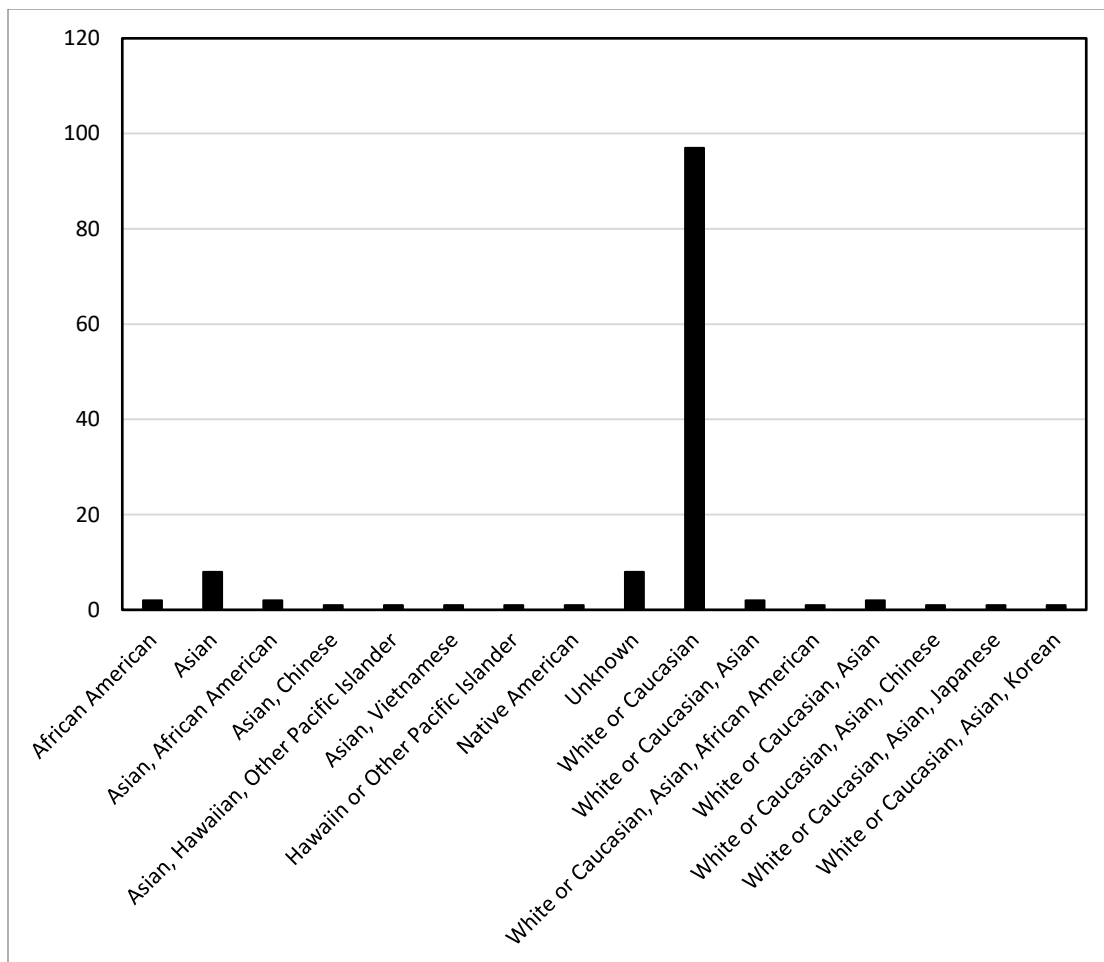


Figure 1. Ethnicity of students enrolled in the school of education, 2016-17.

Sampling Procedures

The longitudinal data were gathered from the 2013-14 school year through the 2016-17 school year. The 2013-14 school year was the first year that edTPA data became consequential for candidates. During this span, a total of 131 participants took the edTPA during the 2013-14 school year, 147 during the 2014-15 school year, 141 during the 2015-16 school year, and 146 during the 2016-17 school year. Participants took a variety of edTPA's content specializations during this period as is represented in Table 1.

Table 1

Number of edTPA Content Area Tests Administered by Year

	Classical Languages	Elementary Literacy	Elementary Math	English as an Additional Language	Family and Consumer Sciences	Health Education	K-12 Performing Arts	K-12 PE	Library Specialist	Secondary English Language Arts	Secondary History Social Studies	Secondary Math	Secondary Science	Special Education	World Languages	Visual Arts
2013-14	0	18	26	2	2	0	7	4	0	8	7	10	17	28	0	2
2014-15	0	21	38	0	0	1	6	2	0	12	10	7	16	29	4	1
2015-16	2	11	32	0	0	1	9	4	2	18	8	9	21	20	4	0
2016-17	0	17	35	1	1	3	8	5	2	10	15	11	11	22	3	2

Candidates participated in both traditional and alternative route programs during this span. Undergraduates (UG) are students who are participating in a traditional undergraduate program at the university and are pursuing a bachelor's degree and Washington State teaching certification (Seattle Pacific University, 2018). Students in this program participate in a broad range of liberal arts courses and select a specialization area for certification. Most field placements are one year long and occur in the final year of the program. Students increase their level of independence and participation in the classroom from one day per week at the onset of the year to multiple days per week as the year progresses (2018). The sum of time spent in this field experience model equates to 20 full weeks of student teaching (2018).

Master of Arts in Teaching (MAT) students possess an accredited bachelor's degree and can earn a master's degree and teaching certification in this program (Seattle Pacific University, 2018). The program typically requires seven quarters to complete approximately 63 credits. A 20-hour autumn field experience is required at the onset of the program, and a full-time field experience is required beginning in February of the final year of the program. The sum of these field experiences equates to approximately 70 days of teaching. The program is designed for working professionals and courses are offered in the evening and online (2018).

The Accelerated Master of Arts in Teaching (AMAT) requires that candidates hold an accredited bachelor's degree (Seattle Pacific University, 2018). Participants begin taking coursework during the summer and continue with courses for the remainder of the academic year while completing a 180-day internship experience. The program requires five quarters to complete. Courses are held in the evening or online.

Participants earn 60 credits and are eligible to earn a master's degree and Washington State teaching certification at the conclusion of the program. An online version (AMAT Online) of this program designed to serve those living in rural areas is also available (2018).

While similar to the university's other teacher certification and master's degree programs, the Accelerated Master in Teaching Math and Science program (AMTMS) provides candidates with courses specifically tailored to instructional practices suited for mathematics and science content, rather than general curricula applicable to other subject-area endorsements (Seattle Pacific University, 2018). An online version of this program (AMTMS Online) designed to serve participants in rural areas and those that cannot conveniently reach campus is also available (2018). Both versions require five quarters to complete and a 180-day teaching experience.

The Professional Educator Standards Board (2018) provides a competitive grant program in which eligible public school districts and higher education institutions can apply for funds designed to offset student costs for participants interested in earning teaching certification in areas deemed as shortage areas such as special education, mathematics, and science. The university that served as the setting for this study participates in this program (Seattle Pacific University, 2018). The program is called the Alternative Route to Certification program for School Employees (ARC-SE). The program requires online and on-campus coursework and completion of a 180-day field experience. A total of 45 credits are required (5000 and 6000 level courses), and the program takes four quarters to complete. Participants may elect to take additional coursework to earn a master's degree (2018).

The Revised Code of Washington (2018) charges the PESB and the State Board for Community and Technical Colleges to exercise authority regarding program approval for teacher preparation and certification programs. Washington State Administrative Codes (2018) provide further specificity regarding development and oversight of teacher preparation programs, including alternative route programs. PESB is accountable to both sets of regulations as they define alternative routes to certification programs. PESB (2018) has identified four routes to certification as alternative: route one for paraeducators with Associate's degrees, route two for classified staff with Bachelor's degrees, route three for "career changers" with Bachelor's degrees, and route four for district staff with Bachelor's degrees employed on conditional or emergency substitute certificates. The university that served as the setting for this study offers a variety of programs, some of which meet PESB's definition of "alternative."

For the purposes of this study, a working definition provided by Mitchell and Romero (2010) will serve to delineate traditional from alternative programs. A traditional program is one requiring coursework and a student teaching experience prior to becoming fully certified and obtaining a bachelor's degree. An alternative program includes a residency, coursework (online, in person, or hybrid), and may offer financial incentives toward certification with or without the opportunity to earn an advanced degree. Most of these programs require applicants to possess a bachelor's degree from an accredited program (2010). The university utilized for this study provides a traditional program (UG) and a variety of alternative route programs; all other programs described above. While not specific to PESB's definitions of alternative route programs, Mitchell

and Romero's (2010) definition is utilized in research in this area and allows comparisons of findings between studies assuming limitations regarding generalization are disclosed.

Sampling Procedures

A convenience sample is a sample of subjects utilized because they are available to the researcher, and not because they are necessarily representative of a larger group (Vogt & Johnson, 2011). All teacher education program participants who took the edTPA from 2013-2017 ($n = 565$) were included in the study as data were readily available and helped to satisfy statistical assumptions described below.

Instrumentation

Researchers at Stanford University developed the edTPA, and it was further refined in cooperation with the American Association for Colleges of Teacher Education (AACTE) and Evaluation Systems (Goldhaber et al., 2017). Evaluation Systems is a member organization of Pearson Evaluation Group, and Pearson oversees the administration of the assessment at this time (2017).

The edTPA is a subject-specific performance assessment that evaluates a common set of teaching principles and instructional strategies that are focused on specific content learning outcomes for P-12 students (Peccheone et al., 2017). There are 27 versions of the edTPA specific to different content specializations (Goldhaber et al., 2017). Each version includes 15 different rubrics; each is scored from 1-5 and weighted equally. The sum of scores may range from 15-75 (2017). Washington State groups the rubrics into three categories: planning, instruction, and assessment. Three additional rubrics (student voice) are required in Washington State but are not included in the candidate's score (2017).

Scores are reported as “pass/fail” based upon a cut score (Pecheone et al., 2017). Pearson provides recommendations as to cut scores, but states have autonomy to adjust cut scores as they see fit (2017).

Most candidates take the edTPA during the field experience required by their teacher preparatory program. The edTPA requires that candidates plan a learning sequence consisting of 3-5 consecutive lessons or in a single 3-5 hour instructional block (Goldhaber et al., 2017). Candidates must prepare instructional materials, record unedited video of instruction, and provide feedback for specific students. Each of the preceding elements is submitted to Pearson electronically for evaluation (2017).

Content validity of the edTPA was established by expert review and confirmatory job analysis (Pecheone et al., 2017). Construct validity was established through Exploratory Factor Analysis and is affirmed annually through Confirmatory Factor Analysis and a Partial Credit Item Response Theory model. Consequential validity is contingent upon how the assessment content and results are referenced in instruction and policy. Studies addressing concurrent validity are lacking, but are beginning to emerge (2017). Predictive validity is mixed in terms of student achievement as measured by MSP data and high in terms of subsequent employment in a public school in Washington State (Goldhaber et al., 2017). Interrater reliability is reported at .887 (Pecheone et al., 2017). Cronbach’s alpha (a measure of internal consistency of raw test scores) reflects the extent to which the items of the assessment measure similar constructs and estimates range from 0-1 (Vogt & Johnson, 2011). Interrater reliability was reported as an overall alpha of .907 (Pecheone et al., 2017).

Chapter Four

Results

This purpose of this study is to ascertain relationships and whether statistically significant differences exist in edTPA scores between candidates who participated in traditional teacher preparation programs or traditional teacher preparation programs. Non-parametric and parametric procedures were used in accordance with the hypotheses tested. Analysis was conducted using IBM SPSS software and Alpha levels of .05 were retained unless noted otherwise.

Description of the Sample

The sample consisted of longitudinal edTPA scores (2013-2017) of students that attended an urban, mid-sized university. Results of subjects who took the test more than once in any academic year were omitted. The investigator eliminated retakes to avoid pretesting as a potential threat to internal validity. Initial attempts were reflected in the final data set. Rubric subscores (numerical) summarized with an overall score of “T” were categorically defined as “pass” for the purposes of the non-parametric analysis. “T” indicated that in the absence of a state cut off score, the score was reviewed by university personnel and was determined to be a passing score. Scores noted as “unable to score” were categorically defined as “fail” for the purpose of the non-parametric analysis. Rubrics 1-15 were used to calculate averages for the planning, instruction, and assessment tasks. Rubrics 16-18 (student voice) were omitted, as they were not consequential to candidates at the time of this study. Teacher preparation programs were categorized as traditional or alternative as defined in Chapter Three.

Statistical Results (Non-Parametric)

The alternative hypothesis (H) tested was that there is a relationship between program type (traditional and alternative route teacher preparation programs) and performance (pass/fail score). The null hypothesis (H_o) tested was that there is no relationship between program type and pass/fail score.

Assumptions for the chi-square test of independence were examined. Data consisted of frequency or counts (Rovai et al., 2014). Data met this assumption.

This procedure requires two variables measured as categories (Rovai et al., 2014). This assumption was satisfied as variables consisted of program type and pass/fail score.

The categories were mutually exclusive; members of one group could not be members of the other (Rovai et al., 2014). Participants either participated in a traditional or alternative program. As such, each subject contributed to only one cell in the contingency table.

The sample size ($N = 550$) exceeded the number of cells multiplied by five (McCugh, 2013).

The chi-square test of independence examined the relationship between teacher preparation program type (traditional or alternative) and edTPA score (pass or fail). Chi-square did not indicate a statistically significant relationship between the variables ($X^2 (1) = .430, p = .05$). Note that differences in expected and actual traditional and alternative program students that passed or failed were small (< 2). The adjusted residual for traditional program and failing score (.7) indicated that there were more failing scores than expected when adjusting for sample size. The adjusted residual for traditional

program and passing score (- .7) indicates that there were fewer passing scores than expected when adjusting for sample size. The adjusted residuals for alternative programs reflect (- .7) for fail and (.7) for pass when adjusted for group size. There was insufficient evidence to reject the null hypothesis. Refer to Table 2 for a summary.

Table 2

Contingency Table Overall edTPA Score

		Fail	Pass	Total	
Program Type	Traditional	Count	8	150	158
		Expected Count	6.6	151.4	158.0
		% of Total	1.5%	27.3%	28.7%
		Adjusted Residual	.7	-.7	
	Alternative	Count	15	377	392
		Expected Count	16.4	375.6	392.0
		% of Total	2.7%	68.5%	71.3%
		Adjusted Residual	-.7	.7	
Total		Count	23	527	550
		Expected Count	23.0	527.0	550.0
		% of Total	4.2%	95.8%	100.0%

Statistical Results (Parametric)

The alternative hypothesis (H) examined with parametric procedures is that there are statistically significant differences in edTPA scores on planning, instruction, or assessment based on program type. The null hypothesis (H_o) is that there are no statistically significant differences in edTPA scores on planning, instruction, or assessment based on program type.

Multivariate Analysis of Variance (MANOVA) requires a number of assumptions. Many are consistent with other parametric procedures, and some are of particular importance when using this procedure (Tabachnick & Fidell, 2013). Data were evaluated to test these assumptions, and the results follow.

The procedure requires the presence of two or more dependent variables measured as intervals or ratios (Tabachnick & Fidell, 2013). The study included three dependent variables that met this requirement. The variables were planning average (PA), instruction average (IA), and assessment average (AA). These are numerical scores, and this assumption was satisfied.

The assumption of sample size ($N = 540$) was satisfied. There were more cases in each group than the number of dependent variables. Ideally, cell sizes are approximately equal, although this is not a requirement (Rovai et al., 2014).

Independent variables must consist of two or more categorical groups (Tabachnick & Fidell, 2013). The study included two categorical groups. The groups consisted of those participating in traditional teacher preparation programs and those participating in alternative route preparation programs.

MANOVA requires independence of observations, that is measurements of one group must not relate to measurements of the other (Tabachnick & Fidell, 2013). Data met this requirement as participants were coded as either members of a traditional teacher education program or an alternative route program as defined in Chapter Three.

MANOVA is sensitive to the presence of outliers (Tabachnick & Fidell, 2013). Conducting a linear regression and analysis of the Mahalbonis distance tested this assumption. The Mahalbonis test is the most common test for the presence of multivariate outliers (Vogt & Johnson, 2011). The Mahalbonis test measures the distance of individual cases from the average of the predictor variables (Field, 2013). Although it is difficult to determine an exact point at which to eliminate cases identified as outliers, Barnett and Lewis (1978) produced a table of critical values dependent upon the number

of predictors and sample size, from which Mahalabonis distances are derived (Field, 2013). The maximum critical value indicated for three dependent variables is 16.27. Eight cases with values ranging from 16.62-68.99 were eliminated.

Gall, Gall, and Borg (2007) reported that researchers often skip testing this assumption [multivariate normality] because MANOVA is robust. Equality of group dispersions can be violated without violating the overall validity of the test. Nevertheless, this assumption was evaluated. Multivariate normality was analyzed with the Shapiro Wilks test. The Shapiro Wilks statistic tests whether sample data have been drawn from a population with a normal distribution, and it can be used with larger sample sizes (Rovai et al., 2014). *P* values greater than .05 indicate normality. PA, IA, and AA = .001. Visual inspection of Normal Q-Q plots approximated straight lines, and this was evidence of normal distribution (Vogt & Johnson, 2011). Values of skewness and kurtosis were + or – 1.00 except for IA which indicated kurtosis at 1.091. Visual examination of histograms provided additional evidence suggesting normal distribution. Given the sensitivity of the Shapiro Wilks test in conjunction with all other evidence, this assumption was satisfied.

Box's M tests the null hypothesis that the observed covariance of the dependent variables are equal across groups and it is considered a very conservative (i.e., sensitive) test (Rovai et al., 2014). Group sizes were unequal, so Box's M could not be disregarded in this case (2014). Box's M is significant at *p* values less than .001, for these data the homogeneity of covariance assumption was satisfied.

A linear relationship between each pair of variables was evaluated utilizing a matrix scatterplot, which is an efficient manner of evaluating this assumption (Rovai et

al., 2014). Visual inspection revealed linear relationships in all pairings and this assumption was satisfied.

MANOVA requires the absence of multicollinearity (Tabachnick & Fidell, 2013). Pearson correlations ranged from .62-.64. Multicollinearity between dependent variables was not indicated, and this assumption was satisfied.

Following a review of all required assumptions, the investigator reaffirmed that all were satisfied and that analysis of data utilizing the MANOVA procedure was appropriate.

The MANOVA was conducted to test the null hypothesis that there were no statistically significant differences in edTPA score (PA, IA, AA) based upon program type (traditional or alternative route). See Table 3 for descriptive statistics.

Table 3

Descriptive Statistics from MANOVA

	Type of Program	Mean	Std. Deviation	N
Planning Average	Undergraduate	3.158	.4409	156
	Alternative	3.249	.5292	385
	Total	3.223	.5066	541
Instructional Average	Undergraduate	3.105	.4484	156
	Alternative	3.228	.5085	385
	Total	3.193	.4947	541
Assessment Average	Undergraduate	3.063	.5224	156
	Alternative	3.193	.5909	385
	Total	3.156	.5745	541

The MANOVA indicated a significant omnibus effect (Pillai's trace = .50). Pillai's trace is highly robust and is used when assumptions of normality may be violated (Vogt & Johnson, 2011). While the significant omnibus effect indicated that differences existed, it did not identify where the specific differences existed. Additional post hoc testing was required.

Post hoc testing to determine where specific statistically significant differences existed was conducted. Partial ETA Squared values for PA were ($p = .057$, $\eta_p^2 .007$) and were not significant. Partial ETA Squared values for IA were ($p = .008$, $\eta_p^2 .013$) and AA were ($p = .017$, $\eta_p^2 .011$). IA variance accounted for was 1.3% and AA variance accounted was 1.1%. Although differences did exist, note that effect sizes were small. Only 1.3% of the variance in IA can be explained by program type, and 1.1% of the variance in AA was explained by program type. Further interpretation of these results follows in Chapter 5.

Summary

Both non-parametric and parametric procedures were used in this study in accordance with the hypotheses tested and characteristics of the data. Although non-parametric procedures are considered less powerful than parametric procedures (Field, 2013), the inclusion of these procedures was appropriate in this case as edTPA data are interpreted by consumers and researchers in many ways.

The results of the chi-square test of independence addressed the question of a relationship between program type (traditional or alternative) and a pass/fail score on the edTPA. Framing the hypothesis in this way reflects the manner in which teacher

preparation candidates speak to their scores (i.e., did you pass or fail) and with how scores are reported to PESB (Professional Educator Standards Board, 2017).

The MANOVA procedure is considered more robust as it requires that data meet assumptions beyond those required for non-parametric procedures (2017). The MANOVA indicated a significant omnibus effect (Pillai's trace = .50). This result can be interpreted to mean that while differences exist, the specific differences cannot be identified without post hoc testing. Post hoc testing to determine where specific differences existed was conducted. Only 1.3% of the variance in IA can be explained by program type, and 1.1% of the variance in AA was explained by program type. Further interpretation of these results follows in Chapter 5.

Chapter Five

Discussion

In 2017, the ED compiled a list of teacher shortage areas nationally and by state. Overall results indicated that shortages existed in special education, science, and mathematics. Teacher shortages in other content areas also existed yet deficits in special education, science, and mathematics were consistent in early every individual state (2017).

National, state, and local efforts have attempted to ameliorate the impacts of teacher shortages. Many states have developed alternative route to certification programs (in conjunction with teacher preparation programs and school districts). These programs are considered educational innovations by the ED (2004). It is estimated that at least 20% of new teachers entering the teaching profession will do so via an alternative route program (Woods, 2016). Despite growing dependence on alternative route programs, research attempting to compare their effectiveness with traditional programs is in its infancy, and few studies exist (Buchanan et al., 2013).

The utilization of a competency exam as a means to measure a teacher's effectiveness is an indirect means of measuring the quality of teacher preparation programs. The requirement to pass a competency-based exam toward teaching licensure is not new, although application of this practice is cyclical (Blackford et al., 2012). The paucity of research comparing traditional versus alternative route programs concerning performance on the edTPA is particularly surprising given the increased usage of the measure despite ongoing controversy among experts in the field (Hebert, 2017; Parks & Powell, 2015). Given the lack of research comparing edTPA performance for students

from traditional versus alternative route programs, this study is a pilot and serves as an initial inquiry to this question.

The first purpose of the study was to compare observed and expected pass/fail rates (categorical data) on a commercially prepared, widely adopted assessment used to ascertain candidate readiness (edTPA) between traditionally prepared and alternatively prepared candidates. The second purpose was to determine if there were statistically significant differences in (continuous data) scores on the planning, instruction, and assessment tasks between traditionally prepared and alternatively prepared candidates on the edTPA.

Overview and Discussion of Findings

Non-parametric research question one. RQ1: Is there a statistically significant difference between observed and expected pass/fail scores for traditionally and alternatively prepared teachers on the edTPA?

The alternative hypothesis (H) is that there is a relationship between program type (traditional and alternative route teacher preparation programs) and performance (pass/fail score). The null hypothesis (H_o) is that there is no relationship between program type and pass/fail score.

The chi-square test of independence examined the relationship between teacher preparation program type (traditional or alternative) and edTPA score (pass or fail). Chi-square did not indicate a statistically significant relationship between the two variables ($X^2 (1) = .430, p = .05$). Note that differences in expected and actual traditional and alternative program students that passed or failed were small (< 2). The adjusted residual for traditional program and failing score (.7) indicated that there were more failing scores

than expected when adjusting for sample size. The adjusted residual for traditional program and passing score (-.7) indicated that there were fewer passing scores than expected when adjusting for sample size. The adjusted residuals for alternative programs reflected (-.7) for fail and (.7) for pass when adjusted for group size. There was insufficient evidence to reject the null hypothesis.

Parametric research question one. RQ1: Are there statistically significant differences in edTPA planning, instruction, or assessment scores between participants in traditional teacher preparation programs and participants in alternative route teacher preparation programs?

The alternative hypothesis (H) is that there are statistically significant differences in edTPA scores on planning, instruction, or assessment based on program type. The null hypothesis (H_0) is that there are no statistically significant differences in edTPA scores on planning, instruction, or assessment based on program type.

The MANOVA was conducted to test the null hypothesis that there were no statistically significant differences in edTPA scores (PA, IA, AA) based upon program type (traditional or alternative route). The MANOVA indicated a significant omnibus effect (Pillai's trace = .50). Pillai's trace is highly robust and is used when assumptions of normality may be violated (Vogt & Johnson, 2011). While the significant omnibus effect indicated that differences existed, it did not identify if the specific differences occurred in PA, IA, or AA. Additional post hoc testing was required.

Post hoc testing to determine where specific differences existed was conducted to identify specific areas where differences occurred. Differences in IA ($p = .008$, Partial Eta

Squared .013) and AA ($p = .017$, Partial Eta Squared .011) existed although these effect sizes were small.

Summary of Results Discussion

Analysis of categorical (pass/fail) edTPA data was a matter of practicality as consumers of the test data often speak to it in terms of passing or failing the overall test. In addition, PESB collects pass/fail data from universities in this manner (PESB, 2017). Recent research by Goldhaber et al. (2017) included edTPA results reported both categorically and numerically.

Given the results, it is tempting to conclude that there is no relationship between program type (traditional or alternative route) and pass/fail score. This conclusion is an oversimplification of the findings. The chi-square test of independence is less robust than parametric procedures because it requires satisfaction of fewer assumptions than parametric methods (Gall et al., 2007). As such, one can only conclude that there was insufficient evidence to reject the null hypothesis based upon the limitations of this procedure. Further analysis was required to ascertain specific differences and was the rationale for the employ of the subsequent parametric test.

The MANOVA reported a significant omnibus effect (Pillai's trace = .50). This result indicated that statistically significant differences in program type and scores in PA, IA, and AA existed, but did not indicate in which assessment task the differences existed. Post hoc testing revealed differences in IA (.008) and AA (.017), but not in PA (.057). Power analysis is usually conducted prior to statistical analysis to ascertain an appropriate sample size although it may also occur following analysis (Vogt & Johnson, 2011). Observed power for IA was .752 and AA was .665. Although power of .8 is desirable

(Vogt & Johnson, 2011), both values approach this recommendation. Partial Eta Squared values were .013 for IA and .011 for AA. These values indicated that 1.3% of the variance in IA scores and 1.1% of the variance in AA scores could be explained by program type. Moreover, the means of student edTPA scores between traditional and alternative program participants in PA, IA, and AA were small. The mean for traditional programs in PA was 3.15 and alternative programs was 3.24. The mean for traditional programs in IA was 3.10 and 3.22 for alternative programs. Traditional programs had a mean of 3.06 in AA while alternative programs had a mean of 3.19. While the overall omnibus effect was statistically significant and post hoc testing revealed differences in IA and AA, they are not practically significant and should not be the basis for programmatic changes without further research.

Although the MANOVA indicated differences in IA and AA, it is important to note that this test reflects the initial attempts at the edTPA for each participant. Subsequent attempts (which the student may have passed) were eliminated to avoid pre-post-test threats to internal validity. Moreover, the vast majority of students in this study passed the edTPA on the first attempt. This may lead the reader to conclude that this study was inconsequential. The investigator reminds the reader that while Washington State's criteria to pass in each of the 23 content specializations have increased through policy revisions, the current score required to pass in most areas (40) is less than the overall national mean in 22 of 23 teaching fields (Pecheone et al., 2017). If Washington State continues to adopt scores commensurate with Pearson's recommendations, it will become increasingly important to attend to overall performance for all students regardless of program type, as well as the performance of students participating in traditional or

alternative programs. The addition of student voice rubrics as consequential will further the need for ongoing analysis of overall results and results by program type.

Limitations and Delimitations of the Study

The first limitation that must be acknowledged is time. McMillan (2012) described this limitation as a recognition that some explanations may change over years or decades. The investigator acknowledges time as a limitation in the review of research comparing traditional and alternative route programs, particularly in their performance on competency-based exams (Buchanan et al., 2013). Further research utilizing edTPA data to compare program types is limited at this time but may emerge in the future.

Although the data met assumptions for each of the statistical procedures utilized, limitations to this study exist and must be acknowledged. The second limitation is that of generalizability. Vogt and Johnson (2011) defined generalizability as the degree to which you can come to conclusions about a population based on a particular sample. Data gathered from a single, mid-sized urban university was utilized. Although the sample utilized in the study was large and helped to satisfy the assumptions of the statistical procedures used, it consisted of scores from a single locale and was a convenience sample. This limits generalizability.

Delimitations are choices deliberately made by the researcher that must be addressed as further limitations to the study. The first such limitation the investigator wishes to disclose was the definition employed to distinguish between traditional and alternative route programs. As previously disclosed, the ED (2004) has identified ten definitions of alternative route programs. Research reviewed in this study often includes some descriptive information of the programs discussed, although programs are usually

broadly classified as traditional or alternative. The investigator utilized a working definition provided by Mitchell and Romero (2010). A traditional program is one requiring coursework and a student teaching experience before becoming certified and obtaining a bachelor's degree. An alternative program includes a residency, coursework (online, in person, or hybrid), and may offer financial incentives toward certification with or without the opportunity to earn an advanced degree. Most alternative programs require applicants to possess an existing bachelor's degree upon entry to the program (2010). While this broad definition reflects the available research comparing programs, the investigator utilized it to classify programs in this particular study. As a result, the findings do not represent differences between distinct alternative route programs available at the university although descriptions of these programs were provided.

Recommendations for Future Research

Recommendations for teacher preparation programs. Lack of consistency in how programs are designed and described makes comparisons of efficacy between programs very difficult. Zigmond (2003) summarizes the challenges inherent in conducting research comparing (special education) programs in the following passage:

Of course, research on the efficacy of special education placements is very hard to conduct at all, let alone to conduct well. For example, definitions of service delivery models or settings vary from researcher to researcher, and descriptions of treatments being implemented in those models or settings are woefully inadequate. Random assignment of students to treatments is seldom an option, and appropriately matched (sufficiently alike) samples of experimental and control students and teachers are rare. As a result, where special education occurs

is not a phenomenon that lends itself to precise investigation, and funding for research studies and publication of results in refereed journals are difficult to achieve. (p. 196)

The challenge identified by Zigmond (2003) is just as relevant when applied to the challenges faced by researchers that have attempted to compare traditional and alternative route programs. Further efforts to delineate specific program features consistently will help researchers to analyze and draw meaningful conclusions about the efficacy of particular alternative programs when compared with each other as opposed to a single class described as “anything other than traditional.” Grossman and Lieb (2010) suggest four elements that may serve as essential markers in alternative route program design. Features of the provider (university, school district, or other program) are easily gathered and may help to discern program similarities and differences. Specific labor markets targeted for recruitment is also an important element. Does the program seek to identify candidates to serve in high needs certification areas or to serve in a specific rural or urban area? Coursework requirements vary greatly among alternative programs but are expressly stated for applicants. As such, this information could be gathered for program comparisons. Finally, eligibility criteria for program entry could be described and used to build an understanding of common and unique aspects of specific alternative programs as compared to traditional programs (2010).

In 2005, the American Educational Research Association (AERA) released a report addressing two primary themes (Cochran-Smith, 2005). AERA’s priority was to present an objective summary of research conducted on the impacts that educational policy has had on preservice education in the United States. A secondary priority was to

recommend a research agenda that would address shortcomings in existing research in teacher preparation to move the field forward. In discussing the second theme, the report stated (2005):

The research comparing the impact of different types of teacher education and programs and pathways (4-year to 5-year, traditional-alternative routes) does not point to the superiority of any one path. However, across the research, there is evidence that certain program components and characteristics are related to teacher quality and pupils' achievement, such as consistent vision, strong communication between universities and schools, certain coursework and school/community fieldwork, and effective use of certain teacher education strategies. (p. 302)

Humphrey and Wechsler (2007) expressed a similar argument. The authors conducted a qualitative study designed to identify characteristics of participants in alternative route programs as well as the programs themselves. Seven alternative route programs were selected, and case studies with participants in these programs included data gathered from interviews, observations, and examination of relevant program documents. Humphrey and Wechsler (2007) summarized their findings in the following, "we find that both sides of the debate [proponents and opponents of alternative route programs] fail to capture the variation in participants' characteristics and experiences in the programs" (p. 483). The authors further concluded that comparing program types was not especially useful and that a better approach would be to study individuals with similar backgrounds, school experiences, and learning opportunities (2007). Humphrey and Wechsler (2007) and Cochran-Smith (2005) call for additional research in teacher

education programs but challenge us to look beyond comparisons of program types toward a deeper analysis of program characteristics (regardless of program type, traditional or alternative). The authors argued for the identification and incorporation of a set of evidence-based practices for use in a variety of program models. Research directed toward these “second order” program characteristics is recommended.

A final recommendation directed toward teacher preparation programs may seem contrary to preceding recommendations for research designed to identify effective practices as opposed to merely comparing programs by type. Hebert (2005) reported that Pearson plans to create local scoring options that will allow university personnel to participate in the analysis of student edTPA assessments. While Hebert expressed salient arguments against aspects of the edTPA assessment, the investigator recommends that university staff in states where a passing score on the edTPA is required for state certification participate in any such opportunity that develops. While many in the field may share Hebert’s concerns, the investigator asserts that our obligation to students and their success overrides philosophical debate on “teaching to the test” or “mandated testing for external accountability.” It is sufficient to say that the test is required and the stakes are high; we must support students.

Recommendations for policymakers. Further research to address the utilization of edTPA as a requirement for teacher certification is recommended. Hebert (2017) cited evidence of a conflict of interest in that much of the research directly related to the edTPA was written by the authors of the edTPA themselves. A review of “supporting research” included on Pearson’s website revealed an excess of 200 articles which were organized by the subtopics of the edTPA but did not include any explanation as to how

the 200 articles related to the edTPA (2017). Moreover, much of the available research conducted on the tool draws comparisons to the PACT assessment (often considered edTPA's precursor), which differs considerably from edTPA in both design and scoring procedures (2017).

Goldhaber et al. (2017) provided an analysis of pass/fail rates among minority candidates. Hispanic candidates in Washington State were more than three times more likely to fail the edTPA when compared with non-Hispanic White candidates. Although the specific reasons for this difference cannot be derived from their study, this investigator concurs with their recommendation that further investigation is recommended (2017). Utilizing a test that may serve to hinder further efforts to diversify the teaching workforce seems contrary to Washington State's efforts to encourage diverse applicants to serve as educators.

The Professional Educator Standards Board (2018) reports institutional scores on edTPA as weighted averages and passing rates. The weighted averages indicate differences between institutions; some fall below the state's required mean and others fall above the mean. From a practical standpoint, it seems worthwhile to investigate why these differences exist. In contrast, the passing rate reported for each university is 100% (PESB, 2018). This percentage begs the question, if this test is a measure of teaching quality and a requirement for certification, is it serving to discriminate between candidates with higher or lower performance when nearly everyone (eventually) seems to pass? Is it worthy of the financial cost to candidates? Given the considerable time required to complete the assessment and the cost to teacher candidates, research

investigating other tools of “competency” that demand less time and cost is warranted if the practice of requiring such a test is to continue.

References

- Aragon, S. (2016). *Teacher shortages: What we know*. Retrieved from <http://www.ecs.org/ec-content/uploads/Teacher-Shortages-What-We-Know.pdf>
- Basinger, J. (2000). Colleges widen alternate routes to teacher certification. *Chronicle of Higher Education*, 18.
- Blackford, K., Olmstead, G., & Stegman, C. (2012). Teacher licensure and qualified teachers: Are certification examinations enough? *American Journal of Educational Studies*, 5(1), 5-19.
- Boers, D. (2007). *History of American education primer*. New York: Peter Lang.
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *American Education Finance*, 176-216.
- Brown, E. (2005). Service-learning in a one-year alternative route to teacher certification: A powerful multicultural teaching tool. *Equity & Excellence in Education*, 38(1), 61-74. doi:10.1080/10665680590907459
- Buchanan, T., Lang, N., & Morin, L. (2013). *Perception of novice teachers from alternative and traditional licensing programs* (Doctoral dissertation, Lipscomb University). Retrieved from <https://www.lipscomb.edu/education/upload/file/62665/perception%20of%20preparedness%20of%20novice%20teachers%20from%20alternative%20and%20traditional%20licensing.pdf>

- Burbank, M., Bates, A., & Schrum, L. (2009). Expanding teacher preparation pathways for paraprofessionals: A recruiting seminar series. *Teacher Education Quarterly*, 199-216.
- Chopin, L. (2013). Untangling public school governance: A proposal to end meaningless federal reform and streamline control in state education agencies. *Loyola Law Review*, 59(2), 399-462.
- Clark, M., Chiang, H., Silva, T., McConnell, S., Sonnenfeld, K., Erbe, A., & Puma, M. (2013). *The effectiveness of secondary math teachers from Teach for America and the Teaching Fellows programs*. Retrieved from National Center for Education Evaluation and Regional Assistance Institute of Education Sciences website: <https://ies.ed.gov/ncee/pubs/20134015/pdf/20134015.pdf>
- Cochran-Smith, M. (2005). Studying teacher education what we know and need to know. *Journal of Teacher Education*, 56(4), 301-306. doi:10.1177/00224871052801116
- Field, A. (2013). *Discovering statistics using IBM SPSS Statistics*. Los Angeles: Sage.
- Fowler, F. (2009). *Survey research methods* (4th ed.). Los Angeles, CA: Sage.
- Fox, A., & Peters, M. (2013). First year teachers: Certification program and assigned subject on their self-efficacy. *Current Issues in Education*, 16(1), 1-16.
- Gall, M., Gall, J., & Borg, W. (2007). *Educational research an introduction*. Boston: Pearson.
- Goldhaber, D., Cowan, J., & Theobald, R. (2017). Evaluating prospective teachers: Testing the predictive validity of the edTPA. *Journal of Teacher Education*, 68(4), 377-393. doi:10.1177/0022487117702582

- Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes. *Journal of Policy Analysis and Management*, 25(1), 75-96.
- Grissom, J. (2008). But do they stay? Addressing issues of teacher retention through alternative certification. In P. Grossman & S. Loeb (Eds.), *Alternative routes to teaching* (pp. 129-155). Cambridge, Massachusetts: Harvard Education Press.
- Grossman, P., & Lieb, S. (2010, May). Learning from multiple routes. *Educational Leadership*. 22-27.
- Hattie, J. (2012). *Visible learning for teachers maximizing impact on learning*. London: Routledge.
- Hebert, C. (2017). What do we really know about edTPA? Research, PACT, and packaging a local teacher performance assessment for national use. *The Education Forum*, 81(1), 68-82. doi:10.1080/00131725.2016.124680
- Hirsch, E. D. (1987). *Cultural literacy: What every American needs to know*. New York: Houghton Mifflin.
- Humphrey, D., & Wechsler, M. (2007). Insights into alternative certification: Initial findings from a national study. *Teachers College Record*, 109(3), 483-530.
- Hussar, W., & Baily, T. (2016). *Projections of education statistics to 2023* (NCES 2015-073). Washington, DC: Government Printing Office.
- Johnson, S., & Birkeland, S. (2003). Pursuing a “sense of success”: New teachers explain their career decisions. *American Education Research Journal*, 40(3), 581-617.

- Knight, S., Lloyd, G., Arbaugh, F., Edmondson, J., Nolan, J., McDonald, S., & Whitney, A. (2012). Getting our own house in order: From brick makers to builders. *Journal of Teacher Education, 63*(1), 5-9. doi:10.1177/0022487111424891
- Kroeger, L., & Ke Vares, K. (2004). *Wisconsin edTPA fast facts*. Retrieved from www.uwosh.edu/coehs/teachered/edtpa/docs/wi-edtpa-fast-facts-loris-version.pdf
- Lowe, D. (2012). *Teacher perception of their initial traditional or alternative teacher training program* (Doctoral dissertation). Retrieved from The Aquila Digital Community.
- McHugh, M. (2013). The Chi-square test of independence. *Biochemia Medica, 23*(2), 143-149. doi:10.11613/BM.2013.018
- McMillan, J. (2012). *Educational research fundamentals for the consumer*. Boston: Pearson.
- Mencken, H. (n.d.). Quotes. Retrieved from https://www.brainyquote.com/quotes/h_1_mencken_129796
- Mitchell, D., & Romero, L. (2010). The politics and practice of teacher certification. *Educational Administration Quarterly, 46*(3), 363-394. doi:10.1177/0013161X10375272
- Parkes, K., & Powell, S. (2015). Is the edTPA the right choice for evaluating teacher readiness? *Arts Education Policy Review, 116*, 103-113. doi:10.1080/10632913.2014.944964
- Pecheone, R., Whittaker, A., & Klesch, H. (2017). *Educational assessment & meaningful support 2016 edTPA administrative report*. Retrieved from https://secure.aacte.org/apps/rl/res_get.php?fid=3621&ref=rl

- Professional Educator Standards Board. (2017). *2017 PESB annual report guidance for the fall of 2017 reporting period*. Retrieved from <https://docs.google.com/document/d/1HFtBn2Sw-PkFMSP33bRtxL6YJss7SsVqJgnfWGTSZK8/edit#>
- Professional Educator Standards Board. (2018). *edTPA scores by program and subject*. Retrieved from <https://www.pesb.wa.gov/data-reports/annual-report-to-the-board/knowledge-skills-assessment/edtpa-scores-by-program-and-subject/>
- Ramsey, P. (2014). Toiling together for social cohesion; International influences on the development of teacher education in the United States. *Paedagogica Historica*, 50(1-2), 109-122. doi:10.1080/00309230.2013.872681
- Revised Code of Washington. (2018). RCW 28A.410.292 *Articulated pathway for teacher preparation and certification: Program approval*. Retrieved from <http://app.leg.wa.gov/RCW/default.aspx?cite=28A.410.292>
- Rovai, A., Baker, J., & Ponton, M. (2014). *Social science research design and statistics; A practitioner's guide to research methods and IBM SPSS Analysis*. Chesapeake, VA: Watertree Press LLC.
- Sass, T. (2013). Licensure and worker quality: A comparison of alternative routes to teaching. Andrew Young School of Policy Studies Research Paper Series, 1-40.
- Seattle Pacific University. (2018a). *School of education, education programs*. Retrieved from <http://spu.edu/academics/school-of-education/education-programs>
- Seattle Pacific University. (2017). *Seattle Pacific University Institutional Review Board users guide*. Retrieved from <http://spu.edu/~media/university-leadership/provost/documents/IRB%20User%20Guidelines%20120317.ashx>

- Seattle Pacific University. (2018b). *Teacher education 2014-2017 demographic data* [Data file]. Retrieved from (link unavailable as data set includes confidential student data)
- State of Washington Professional Educator Standards Board. (2018c). *edTPA policies*. Retrieved from <https://www.pesb.wa.gov/preparation-programs/assessments/performance-assessments-edtpa/edtpa-policies/>
- State of Washington Professional Educator Standards Board. (2018a). *Alternative routes to teacher certification*. Retrieved from <http://apps.leg.wa.gov/wac/default.aspx?cite=181-78A>
- State of Washington Professional Educator Standards Board. (2018b). *Alternative routes to teaching block grants*. Retrieved from <https://www.pesb.wa.gov/innovation-policy/grants-pilots/alternative-route-to-teaching-block-grants/>
- Strauss, S. (1993). Theories of learning and development for academics and educators. *Educational Psychologist*, 28(3), 191-203. doi:10.1207/s15326985ep2803_1
- Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the U.S.* Palo Alto, CA: Learning Policy Institute.
- Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson.
- Tournaki, N., Lyublinskaya, I., & Carolan, B. (2009). Pathways to teacher certification: Does it really matter when it comes to efficacy and effectiveness? *Action in Teacher Education*, 30(4), 96-109.

- U.S. Department of Education, Office of Innovation and Improvement. (2004). *Alternative routes to teacher certification*. Retrieved from <https://www2.ed.gov/admins/tchrqual/recruit/altroutes/report.pdf>
- U.S. Department of Education, Office of Postsecondary Education. (2017). *Teacher shortage areas nationwide listing 1990-1991 through 2017-2018*. Retrieved from <https://www2.ed.gov/about/offices/list/ope/pol/ateachershortageareasreport2017-18.pdf>
- U.S. Department of Education, Office of State Support. (2015). *Fundamental change innovation in America's schools under Race to the Top*. Retrieved from <https://www2.ed.gov/programs/racetothetop/rttfinalrptexecsumm.pdf>
- Vogt, W., & Johnson, R. (2011). *Dictionary of statistics and methodology a nontechnical guide for the social sciences* (2nd ed.). Los Angeles: Sage.
- Washington Administrative Code. (2018). *Chapter 181-78A WAC Approval standards for performance-based preparation programs for teachers, administrators, and educational staff associates*. Retrieved from <http://apps.leg.wa.gov/wac/default.aspx?cite=181-78A>
- Washington State Professional Educator Standards Board, edTPA Student Voice Committee. (2016). *edTPA student voice committee's recommendations*. Retrieved from https://drive.google.com/file/d/0B2bHo3NTb_bJYmdnb1Z5dzNBWkU/view
- Woods, J. (2016). *Mitigating teacher shortages: Alternative teacher certification*. Retrieved from <http://www.ecs.org/ec-content/uploads/Mitigating-Teacher-Shortages-Alternative-Certification.pdf>

Zhang, G., & Zeller, N. (2016). A longitudinal investigation of the relationship between teacher preparation and teacher retention. *Teacher Education Quarterly*, 43(2), 73-92.

Zigmond, N. (2003). Where should students with disabilities receive special education services? *The Journal of Special Education*, 37, 193-199.