Aligning Policy to Performance: The edTPA policy and program improvement in Washington State's Teacher Preparation Programs

Maria Muzzo

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Deborah McCutchen, Chair

Charles A. Peck

William Zumeta

Program Authorized to Offer Degree:

Education

University of Washington

Abstract

Aligning Policy to Performance: The edTPA policy and program improvement in Washington State's Teacher Preparation Programs

Maria Muzzo

Chair of the Supervisory Committee:

Deborah McCutchen, Chair

Associate Dean,

College of Education

One stated purpose for the use of the edTPA is to guide teacher preparation program improvement (Stanford Center for Assessment, Learning, and Equity, 2013). The Washington State Legislature included teacher preparation program reform on its agenda, and requested that the Professional Educator Standards Board identify ways in which the edTPA "will be included in state reported data on preparation program quality" (RCW 28A.410.270, 2013). While teacher preparation programs in Washington could state what actions were taken to reform their practices, it is difficult to identify which components of the policy have been effective (Peck, Singer-Gabella, Sloan, & Lin, 2014). This study seeks to identify the status of teacher preparation

program reform in the colleges of education in Washington State that piloted the edTPA for four years. Specifically, this study seeks to identify the extent of change to integrate the *core concepts* measured by the edTPA (planning, instruction, and assessment) at the program and instructional levels, and the extent to which faculty and field personnel attribute the edTPA's core concepts as aligned to their college's mission.

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

ACKNOWLEDGEMENTS

I would like to thank my family and friends for their ongoing support, for cheering me on, and for picking me up when I fell. And to Cathy Taylor, for fighting the good, statistical fight to find truth in the headwinds of politics. And finally, I would like to thank my chair, Deborah McCutchen, for challenging my understandings, and for walking my theories from confusion to articulation with extraordinary patience.

DEDICATION

This dissertation is dedicated to my parents, who walk me through life's challenges every day. I am honored to have learned from you.

To Chuck and Jo – you are models of persistence, intelligence, and lifelong learning. The dissertation is in the mail!

And to Ski, who had to put her dreams of a Ph.D. on hold to battle cancer. Your strength is your accomplishment. This one's for you, palie.

Chapter 1. INTRODUCTION

The call for teacher preparation program reform built momentum at the turn of the 21st century (Wiseman, 2012; Lewis & Young, 2013). This call to action stemmed from multiple sources: the need for new strategies in a complex, technical and media-rich society (Darling-Hammond & McLaughlin, 1999; Mintop & Sunderman, 2009; Darling-Hammond & McLaughlin, 2011, Robb, L., 2013; Cochran-Smith & Villegas, 2015; Organization for Economic Development, 2005); wide achievement gaps between diverse learners in the US (Wiseman, 2012, Ratner & Kolman, 2016; Cochran-Smith & Villegas, 2015); low student achievement scores as compared to other developed nations around the world (Organization for Economic Development, [OECD], 2005); and the teacher preparation programs themselves (Wiseman, 2012; Tatto, 2006). In the United States, national databases became available to study student and teacher performance indicators at local levels (Mintrop & Sunderman, 2009), in an effort to improve student performance. Teacher performance was found to be one source of variation in student scores (Paufler & Amrein-Beardsley, 2016), bringing teacher preparation programs under scrutiny (Wiseman, 2012; Cochran-Smith & Villegas, 2015).

In response to the national conversations around teacher preparation program reform, Washington state's colleges of education began working collaboratively with legislative leadership in 2007, developing teaching standards, and piloting teacher performance assessments. Five years later a federal bill was introduced, calling for reform in teacher preparation programs.

Literature suggests three themes that mediate educational policy implementation: 1) structured, program-wide opportunities help faculty and field personnel make sense of the policy's message in order to analyze how to best improve the teacher preparation program (Peck, Gallucci & Sloan, 2010; Peck, Singer-Gabella, Sloan & Lin, 2014), 2) leadership must address faculty autonomy in this era of standardization, in order to facilitate cohesive interpretation of the policy's message and to facilitate an implementation at the local level that drives fundamental change in the instruction of teacher candidates (Coburn, Hill & Spillane, 2016; Mitescu-Reagan, Schram, McCurdy, Change, & Evans, 2016, Meyer & Rowan, 2006), and 3) key stakeholder valuing of the policy's message at the local (implementer, or faculty) level is critical for the adoption of the policy (Evans, 2015; Weatherly and Lipsky, 1977). The purpose of the current study was to describe implementation of teacher education reform that occurred in Washington state teacher preparation programs following passage of the mandate for a teacher performance assessment and, consequently, the adoption of the edTPA. The edTPA is a pre-service teacher performance assessment that measures teacher candidate performance based on a set of articulated standards. By understanding national and local events in the context of policy implementation theory, we can understand the manner in which (if at all) teacher performance assessment legislation drove change within teacher preparation programs. The present study sought to inform our understanding of implementation of the edTPA through the perceptions of faculty and field personnel, and focused on the perceived changes made within teacher education programs and teacher candidate instruction to accommodate the new reform policy. I begin with a review of literature situating the edTPA in

policy implementation theory. A brief historical context of the development of the edTPA stemming from an era of accountability is then provided, in order to contextualize the policy, and subsequent approaches to edTPA implementation. A discussion follows of the political strategies specific to edTPA implementation that may have impacted edTPA adoption within teacher preparation programs and teacher educator practice. Chapter two concludes with a suggested matrix as a conceptual frame for understanding how alignment of resources and supports can balance accountability measures to drive fundamental change. Chapter three describes the research design and methodology for studying teacher preparation program faculty perception of the status of the edTPA implementation in Washington state. Chapter four provides analysis of the research study findings, and finally, chapter five discusses the implications of the findings.

1.1 Search Strategy for the Literature Review

My search for peer-reviewed research conducted in the US, and published between 2010 and 2017 (with a full-text PDF available), related to teacher education reform and policy implementation theory included EBSCO Host, Education Source, ERIC, Historical Abstracts, Library and Information Science Source, PsychInfo, Teacher Reference Center, and Newspaper Source Plus. The title was read, with only articles that appeared to situate the edTPA in teacher preparation program reform efforts selected. When in doubt, the abstract provided information for inclusion / exclusion decisions. Google scholar was consulted for specific peer-reviewed articles that were available in full text online. References of key articles that were the primary contributions to this paper were reviewed to identify more sources. In order to chronicle

events, a search of Education Week newspaper archives yielded useful results for the keyword, “edTPA”.

The first search conducted used the key words “Teacher Preparation Program” and “Reform”. I was attempting to establish where studies of teacher preparation programs were situated in research historically in terms of the educational reform era. Articles that focused on political and implementation tensions within teacher preparation programs were included. Articles that focused on teacher candidate learning were excluded in order to maintain alignment with the study’s focus on teacher educators and the programs themselves. The search engines noted above were used again to search the key words “Teacher Preparation Program” and “edTPA”. This second search was conducted in order to understand what scholarly conclusions had been drawn regarding edTPA’s impact on teacher preparation programs. Again, articles that focused on political and implementation tensions within teacher preparation programs with specific regard to the edTPA were included. Articles that focused on teacher candidate learning and studies linking the edTPA to specific teaching paths were excluded.

Finally, the phrases “Policy Implementation Theory” and “higher education” were used. This final search solidified my understanding of where policy implementation is currently heading, and how it might be useful to guide research.

1.2 Definition of Terms

Within the context of this discussion, the edTPA policy is RCW (Revised Code of Washington) 28A.410.270 (RCW 28A.410.270, 2013), which called for Washington

state to adopt research-based standards ("the standards") defining quality teaching, a beginning teacher performance assessment (the edTPA was chosen), and the manner in which the assessment scores would be used to report on teacher preparation program quality. The edTPA is a portfolio-based performance assessment for pre-service teachers. Touted as both a formative and summative assessment by its designers, the edTPA was designed to determine beginning teacher readiness to teach and could be used for teacher preparation program improvement purposes (Stanford Center for Assessment, Learning, and Equity (SCALE), 2013). The term *core concepts* refers to the overarching teaching knowledge and skills measured by the edTPA: planning, instruction, and assessment. *Standards* refers to the standard criteria defining desired teaching practices within the core concepts, as stated within the edTPA's rubrics. *Rubrics* refer to the standards as written in a measurable format. Finally, the term *passing standard* refers to the cut score, or score designated as required in order to pass the edTPA. The term *environmental infrastructure* refers to the teacher preparation program's policies, procedures and faculty development opportunities in place in the environment where the implementation takes place, which become factors of influence on implementer adoption.

Chapter 2. LITERATURE REVIEW

2.1 THEORETICAL FRAMEWORK

Policy implementation theory provides a useful frame for understanding the extent to which change occurred within teacher preparation programs in response to policy mandate RCW 28A.410.270 (RCW 28A.410.270, 2013). First, Lane (1987) defined implementation as both a *process* of bringing about outcomes of a policy goal and the *outcome* itself. In other words, implementation could be both an ongoing process of continuous improvement and achievement of that process. Lane's (1987) article prompts the question that if implementation is evolution, if it is by nature change, then how can that change be managed to promote outcomes aligned with the policy's objectives? As educational reform is often focused heavily on top-down policy instruments like accountability, to what extent are local implementers aligning their practice with RCW 28A.410.270's objectives of improving teacher preparation?

Finally, Coburn et al. (2016) provide a conceptual frame for understanding the interactions of accountability and alignment, as well as whether minimal, superficial, uneven, or fundamental change has occurred within an organization. In the case of this study, changes refer to those differences in actions meant to improve the teacher preparation program. The conceptual framework illuminates opportunities for teacher preparation program improvement, illustrating the idea that alignment of policy goals with local processes, such as professional development, administrative support (such as time, opportunities to collaboratively use data to analyze practice), and accountability

instruments such as high-stakes test scores can assist faculty in improving practice (Coburn et al., 2016). While the theory is based on research regarding how educational policies are implemented in schools, I use the conceptual frame to understand teacher preparation program improvement. The conceptual frame is discussed in depth below.

Although the theory itself has evolved through time, policy implementation analysts are concerned with understanding the conditions that influence change, and whether changes that occurred were congruent to the policy's goals (Lane, 1987; Coburn et al., 2016).

Since the 1970s, policy implementation theory has evolved from defining implementation and its primary factors of influence (Lane, 1987, Meyer & Rowan, 2006, Honig, 2006, Kohoutek, 2013) to assuming a hierarchical structure of control during implementations would occur, especially when instruments such as accountability, sanctions, and incentives were used, (Hood, 1976, Sabatier & Mazmanian, 1979; Lane, 1987), to arguing that it was local actors involved in the implementation and their interpretations of the policy that determined the policy's outcome (Weatherly & Lipsky, 1977, Kohoutek, 2013). Subsequently, Sabatier (1986) argued the importance of understanding the nuances in implementations, such as the influence of local actors and the manner in which they interpreted the policy's goals during implementations. In the 1980's Sabatier (1986) delineated an Advocacy Coalition model of policy implementation theory, to acknowledge the influence of local stakeholders in policy implementation, while simultaneously reiterating the power of instruments like accountability in implementations.

An outcome of the 1980's standards-based reform movement in education was the creation of academic standards benchmarking the knowledge and skills students in kindergarten through grade twelve (K-12) should know and be able to demonstrate. There was an attempt to align K-12 academic standards with local accountability policies (Darling Hammond, 1990), through student assessments or requirements for teacher preparation, as well as with mandated professional development, and curricular materials (Coburn et al., 2016, Kohoutek, 2013).

Research on teacher practice (Coburn et al., 2016; Weatherly & Lipsky, 1977; Spillane, Reiser & gomez, 2006; Honig, 2006) theorized that in-service teachers who did not demonstrate fundamental change in their teaching practices had developed their own idiosyncratic understandings of policies based on prior knowledge and had been provided little support to understand and respond to the local policies. For example, the lack of alignment between accountability policy and in-service teacher professional development resulted, in some cases, in what some researches call "superficial implementation", such as using story problems as a proxy for teaching authentic problem solving in mathematics (Coburn, et al., 2016). In other words, some teachers may have been using the same pedagogical strategies they had always implemented, rather than fundamentally changing their practice to incorporate the new accountability standards (Coburn, et al., 2016). This "superficial implementation" may have been the first step in interpreting what the policy requirements meant at the local level, to the individual instructor, prior to fundamental, collective innovations (Peck et al., 2009). Researchers argued that structured opportunities using assessment data and collaboration to collectively investigate current practices could help in-service teachers

and faculty fundamentally improve their practice (Peck et al., 2009; Peck & McDonald, 2013; Bastian, Lys, & Pan, 2017). Additionally, multiple initiatives with curricular materials, tests, and often professional development provided as support could have contributed to misinterpretation of the policy's message, and led to divergent demands on teachers, who were left to their own, inevitably variable interpretations of the reform efforts (Cochran-Smith & Villegas, 2015).

Educational policy implementation theorists emphasized the importance of using multiple, appropriate lenses at different intervals of policy processes (Sabatier, 1986; Weible, Heikkila, DeLeon & Sabatier, 2012). Educational policy implementation studies documented that policy implementation outcomes depend on how leadership supports the implementation and holds implementers accountable, how local actors interpret and enact the policy, and how the environmental infrastructure (the place where the policy is implemented and its institutionalized culture of practice) supports use of instruments such as curriculum, tests, and professional development for teachers or faculty (Honig, 2006; Lane, 1987; Coburn, et al., 2016; Meyer & Rowan, 2006). Implementation research stressed the importance of studying the *interaction* between policy, people, and place – the environmental infrastructure consisting of a teacher preparation program's structured opportunities to understand which mediating factors influence outcomes congruent with the goals of the policy (Honig 2006; Coburn, et al., 2016; Durand, Lawson, Wilcox & Schiller, 2016).

During the 2000s, high-stakes accountability practices spurred by the No Child Left Behind (NCLB) Act (2001) led to increased testing and reporting of scores. Classroom scores were published in newspapers to pressure teachers to improve their

practice. Low-performing schools were sanctioned, and those that did not deliver "Adequate Yearly Progress," even for one of many subgroups, were closed or put under new leadership. Although there were slight increases in student achievement, Coburn, et al. (2016) found results of NCLB policy implementation and its associated accountability measures changed the content taught by teachers, but not their pedagogy. Additionally, different policy instruments, such as mandates, high stakes assessments, or performance alignment strategies, that is, the activities a program chooses to engage in to align teacher performance to priorities, were found to differently affect the manner in which educational policies were embraced at the local level (Honig & Rainey, 2015), strengthening the need for analyzing nuances between accountability and alignment. For example, although Coburn, et al. (2016) found that the use of accountability measures led to superficial alignment, Honig and Rainey (2015) found that teaching and learning benefitted when school district leadership and central office supported deeper learning through "performance alignment", structuring priorities in daily practice to facilitate the desired outcome.

National policy calling for accountability measures influenced Washington state's development and implementation of two specific educational policies aimed at improving the instruction of teacher candidates within teacher preparation programs. Washington's policies were embedded in the national educational era of standards and accountability-based reforms previously described (Coburn, et al., 2016) and built on California's experience developing and implementing its teacher performance assessment, the PACT (Pecheone & Chung, 2006). Washington called for teacher educational standards to be developed, a governing board be assigned to guide

implementation of the standards, inform the state of how data would be used to guide preparation program quality (RCW28A.410.210, 2013), and use of a test for teacher candidates (RCW28A.410.270, 2013). A pre-service teacher performance assessment, the edTPA, was developed based on California's PACT (Pecheone & Chung, 2006; Stanford Center for Assessment, Learning, and Equity [SCALE], 2013) and implemented based on lessons learned from its developers (Coburn, et al., 2016; RCW28A.410.210, 2013; RCW28A.410.270, 2013; Peck & McDonald, 2014; Mitescu-Reagan, et al., 2016). PACT was a locally-designed portfolio assessment created by California teacher educators for formative and summative purposes, with tasks culminating into a cohesive teaching event (Pecheone & Chung, 2006). The PACT was found to engage teacher educators in analyzing and improving their practice by: using the assessment formatively and collaboratively to identify student progress, helping analyzing data from PACT to inform faculty of their own strengths and weaknesses in instruction, and maintaining agency among faculty in terms of what and how they taught (Peck et al., 2010).

Developed by a team led by researchers who had worked on the PACT (Pecheone & Chung, 2006), the edTPA was developed as a high-stakes portfolio assessment for teacher candidates and for teacher preparation programs in multiple states across the country (SCALE, 2013; Bastian, Lys, & Pan, 2017; Lewis & Young, 2013). As part of Washington's policy, a teacher candidate's passing score on the edTPA was one of multiple measures required for program completion and ultimately, for teacher licensure (RCW28A.410.270, 2013). The edTPA developers introduced the

exam as a resource with dual purpose, to both assess beginning teacher performance and for teacher preparation program improvement (SCALE, 2013).

The state legislature assigned the Professional Educators Standards Board (PESB) the task of leading the policy's implementation (RCW28A.410.210, 2013). PESB collaborated with the Washington Association for Colleges of Teacher Education (WACTE) to co-develop the teacher performance assessment (edTPA) with Stanford Center for Assessment, Learning and Equity (SCALE), and to implement it statewide. WACTE representatives were comprised of deans and faculty leaders from Washington's teacher preparation programs. The state adopted the edTPA as its performance assessment, piloted the test from 2009 - 2013, and implemented accountability requirements beginning in 2014. Teacher candidate scores from the edTPA were aggregated by teacher preparation program and reported to the state as one measure of preparation program quality (RCW 28A.410.270, 2013).

2.2 Policy Implementation

Both top-down and bottom-up influences were used to drive edTPA adoption throughout the United States. The edTPA's developers used federal accountability measures to build a coalition of support for the instrument from teachers' unions, the National Educational Association (NEA) and the American Federation of Teachers (AFT), and from the American Association for Colleges of Teacher Education (AACTE), before state accountability policies were in place (Hutt, Gottlieb & Cohen, 2018; Coburn, et al., 2016; Kohoutek, 2013). Then the coalitions were used to drive partnership-style edTPA development, piloting, adoption and implementation throughout 29 states and 400 institutions (SCALE, 2017; Hutt, Gottlieb & Cohen,

2018). The edTPA, like PACT (Pecheone & Chung, 2006), was portrayed as an opportunity for coherent learning processes, or "sense-making," both by teacher candidates (Chung, 2008) and by teacher preparation programs (Peck, et al., 2014; Darling-Hammond, 2006a; Peck, Galluci, Sloan & Lippincutt, 2009; Peck & McDonald, 2013; Peck, Muzzo & Sexton, 2010), which could provide a bottom-up force driving alignment of standards of best teacher practices with performance.

The colleges of education in Washington agreed uniformly to use the edTPA to assess teacher candidate performance. A group of universities within the state began piloting the assessment in 2009, while contributing to its development.

Similar to PACT, the edTPA is a portfolio assessment that includes the teacher candidate's lesson plans, a video depicting instruction and student engagement, teacher candidate assessment artifacts such as student quizzes or test feedback, and performance reflections by the candidate. Unlike PACT, teacher candidates were required to submit their edTPA portfolio to Pearson Education, Inc. for evaluation through software provided by their teacher preparation program. The universities ensured technology was in place to link candidate submissions to Pearson Education, Inc., and developed administrative and instructional tools such as handbooks, teacher educator workshops, online faculty support/feedback groups to monitor and support the implementation, and subsequently designated test coordinators. Teacher educator feedback and data from Washington's early pilots informed SCALE, the edTPA's developer, and the American Association of Colleges for Teacher Education (AACTE), the political organization promoting national use of a teacher performance assessment. The coordinated effort to implement edTPA as a national assessment afforded programs

the opportunity to align policy design to program support, and ultimately, hopefully, to teacher candidate instructional performance as state and national policy developed.

2.2.1 *Federal Teacher Performance Assessment Policy*

In 2012, Representative Mike Honda (D-CA) and Senator Jack Reed (D-RI) introduced a bill in Congress endorsed by the American Association of Colleges for Teacher Education (AACTE), to reauthorize and change federal laws regarding teacher preparation (Sawchuk, 2012). Specifically, the bill required teacher preparation programs to report outcome data – academic performance of students taught by teacher candidates, retention of graduates in teacher preparation programs over three years, and results from teacher evaluations of content knowledge, *unless* the programs used a teacher performance assessment. Pass rates from a teacher performance assessment would count instead of the growth measures from student test scores. The bill was endorsed by AACTE, National Association of Elementary School Principals, National Associate of Secondary School Principals, Alliance for Excellent Education, and the National School Boards Association (Sawchuk, 2012; RCW 28A.410.270, 2013).

Arguments for the value of standardized assessments of teacher performance over outcome data as sources of information for program improvement seemed to be persuasive (e.g., Darling-Hammond, 2006a; Peck et al., 2009; Peck et al., 2014), as many states chose to implement teacher performance assessments rather than report outcome data. This appeared to be a critical point for Washington's teacher preparation programs and SCALE to determine how they would implement the assessment, align to the edTPA's standards, and drive value of the assessment throughout the program faculty. Multiple states began implementation of the edTPA with varying amounts of

piloting time, with calls for alignment strategies like faculty development, and curriculum aligned to edTPA (Ratner & Kolman, 2016; Darling-Hammond, 2006b). However the literature highlights multiple tensions occurring at once, which could have impacted the manner in which the edTPA was adopted in various locations. The first tension regarded the shift toward alignment of edTPA standards to teacher candidate instruction – a professionalization approach to teaching (Coburn et al., 2016; Evans, 2015), and away from local control and autonomy (Mitescu-Reagan, et al., 2016). For example, the scoring of the edTPA was found, in some cases, to be demoralizing to program faculty in that it relied on external scorers from Pearson Education, rather than faculty within the teacher preparation programs (Cody, 2012; Mitescu-Reagan et al., 2016). Scoring teacher candidate performance was traditionally the job of the teacher preparation programs. Moving candidate scoring to an outside source was interpreted as a lack of trust in the objectivity of faculty judgment and diminished a key component thought to contribute to the value of teacher performance assessments for program improvement (Peck et al., 2009; Peck et al., 2010; Peck et al., 2014).

Secondly, the literature attributes success of the PACT in California to its ability to engage faculty in improving their instruction through their analysis of teacher candidate performance, and through formative, signature assessments (Mitescu-Reagan, et al., 2016, Peck et al., 2010). However, the edTPA, as a national assessment, outgrew its local roots and grew into a federal policy, which limited local teacher preparation programs' impact in the assessment's development in multiple states. This could have led to faculty perception of diminishing importance of their instruction and expertise. Multiple studies report the criticality of local implementers to be afforded opportunities

to make sense of the policy and value its message (Evans, 2015, Peck et al., 2010), as attitude toward policy has been found to affect the manner in which the policy is implemented. Of six states that Mitescu-Reagan et al. (2016) studied later, three states allowed teacher preparation programs to determine which teacher performance assessment they would use (CA, IA, TN), two states determined with recommendation and commitment from teacher education organizations that the edTPA would be used (MN, WA), and one state immediately enacted the edTPA at the state level without teacher preparation program input, and connected passing rates to Race to the Top funding for teacher preparation programs (NY).

2.2.2 *Washington State – Capacity Building as a Policy Lever*

The federal emphasis on teacher preparation reform provided the impetus for states and teacher preparation programs within those states to better understand the details of how their implementations of teacher performance assessments could be successful and generate useful feedback to programs (Peck et al., 2009; Peck & McDonald, 2013). The Washington Legislature's policy response in the form of two separate laws called for the development of teaching standards, a standards-based assessment to measure and report teacher candidate performance, and inclusion of the assessment as data to inform judgments about program quality (RCW28A.410.210, 2013; RCW28A.410.270, 2013). One bill stated:

> The Professional Educator Standards Board shall adopt a set of articulated teacher knowledge, skill, and performance standards for effective teaching that are evidence-based, measurable, meaningful, and documented in high quality

research as being associated with improved student learning (RCW28A.410.270, 2013).

A second bill provided for a state leadership team to implement the policy and to instruct the state how the assessment "will be included in state-reported data on preparation program quality…[and] how the program produces effective teachers as evidenced by the measures established" (RCW 28A.410.210, 2013). Washington State's Professional Educators Standards Board (PESB) was designated, as the state's leadership team, to coordinate the development of the test and its implementation as a collaborative effort between teacher preparation program leadership involved in the Washington Association for Colleges of Teacher Education (WACTE), Washington State, the test's developers at Stanford University, and the test's promoter, the American Association for Colleges of Teacher Education (AACTE) (SCALE, 2013). As recommended by WACTE, PESB designated the edTPA as its required, pre-service teacher performance assessment, in response to the Washington State Legislature's policy, RCW 28A.410.270. The policy's intent to identify teacher performance is summarized in its statement, "a system that clearly defines, supports, measures, and recognizes effective teaching and leadership is one of the most important investments to be made" (RCW 28A.410.270, 2013).

Washington's teacher preparation programs were in their final year of piloting the edTPA when the state rolled out the assessment (Sawchuk, 2013) and approved the teaching standards (RCW 28A.410.270, 2013; RCW 28A.410.210, 2013). Teacher educators, policy makers, national and regional teacher organizations, as well as teacher

preparation programs, worked together to implement the policies as a means for teacher preparation program improvement (Mitescu-Reagan, et al., 2016).

Washington state employed several strategies to address tensions of autonomy resulting from concerns that the edTPA standards could marginalize other approaches to teacher education, such as multicultural education or social justice. As the edTPA grew from a locally-developed formative and summative assessment to a national assessment designed to professionalize teaching, there were concerns that "alignment" to the edTPA would drive and/or narrow teacher education curricula toward only what could be measured, (Cody, 2012; Mitescu-Reagan et al., 2016). Coburn et al. (2016) describe "alignment" of policy to curricula, program supports, and assessments as a process for driving fundamental change when coupled with high-stakes accountability. However, alignment to the edTPA presented an instructional dilemma for teacher educators – although teacher performance assessments were found to improve consistency of messaging to teacher candidates in terms of what good teaching is (Peck et al, 2010; Mitescu-Reagan et al., 2016), aligning curriculum to edTPA standards could be seen as privileging certain curricula and diminishing or eliminating other, equally important messages to teacher candidates (Mitescu-Reagan et al., 2016). This colonization concern stems from *how* teachers interpreted and valued the concept of aligning curricula to the edTPA's standards – if the standards were imbued as collaborative effort of research-based best practices (Peck et al., 2010), practices that were valued by the teacher educators and a part of their program's goals, then aligning to the standards made sense. However, if aligning to edTPA standards meant changes in the vision and goals of the teacher education program, diluting faculty's autonomy to

decide what should be taught, then alignment represented intellectual colonization, spawning concern that the teacher preparation programs might be taken over by the very instrument designed to evaluate them (Evans, 2015; Mitescu-Reagan, et al., 2016).

It appears that Washington state addressed the tension of curricula colonization versus collaboration by spending four years piloting the assessment, while simultaneously inviting and responding to teacher educator concerns (Honig & Rainey, 2015). This prolonged pilot seemed to encourage local engagement (Mitescu-Reagan, et al., 2016). PESB members held conferences and faculty members held professional development meetings for this purpose both within and across universities. PESB read and responded to faculty feedback through a discussion board during the pilots, in order to build capacity for implementing the edTPA on a tactical level and for identifying instructional adherence to the edTPA's standards. Faculty feedback led, in some cases, to revisions in the edTPA handbooks and in the standards themselves (Pecheone & Whittaker, 2013b).

In comparison to Washington's implementation of the edTPA, instruments used by other states to achieve policy goals of teacher preparation program improvement sometimes foreshadowed the implementation's outcome. There were significant differences across states in the way that the state policy around a teacher performance assessment was implemented, especially with regard to the states' purposes for implementation and stakeholder involvement (Mitescu-Reagan et al., 2016; Hutt et al., 2018). The overarching purpose of the states' teacher performance assessment policies was uniform, calling for a valid and reliable teacher performance assessment to determine the quality of teacher preparation programs through their teacher candidate

pass rates. However, the manner in which states approached the policy differed. The policy was used for program accreditation as a state function in some cases and for program-ranking purposes in other cases (Mitescu-Reagan, et al., 2016). For teacher candidates, the policy was used for state certification requirements, such as to signify teacher preparation program completion, or recommendation for certification (Mitescu-Reagan, et al., 2016).

The requirement to use edTPA grew nationally, despite stakeholder concerns and differing approaches. By October, 2013, twenty-two states and Washington DC were using the edTPA, with eleven more states working toward implementation (Courrege-Casey, 2013; SCALE, 2013). In December, 2013, Washington and New York announced that they would officially implement the edTPA as a high stakes teacher performance assessment required for certification, with five other states intending to use it for part of teacher certification or preparation program improvement by the 2015-16 school year (Sawchuk, 2013a). Washington and New York also set passing standards (Heltin, 2013; Sawchuk, 2013b). SCALE, the organization that developed the edTPA with teacher input (Sato, 2014; Henry, Campbell, Thompson, Patriarca, Luterbach, Lys & Covington, 2013), recommended a maximum total passing standard of 42, with a top score of 75 (Stanford Center for Assessment, Learning, and Equity, 2014), estimating that about 58% of the candidates would pass it on the first try (Heltin, 2013; Sawchuk, 2013c). However, Washington state deferred to a panel of teacher educators, and set the passing standard to 35, for middle and high school teacher candidates (Sawchuk, 2013c).

2.2.3 *Teacher Preparation Program Alignment in Washington State*

Washington State's implementation process included collaboration with stakeholders to develop a sense of agency, as exemplified in its approach to setting the passing standard. By building structured support, opportunities for faculty to make sense of the policy and capacity within educational organizations and the teacher preparation programs themselves, PESB drove the impetus for program improvement from within the programs. Teacher preparation programs developed workshops for collaborative analysis of student work, resulting in considerable change in instruction early in the implementation process (Peck et al., 2009). Programs conceived of these workshops as facilitating a common language of practice, and as driving coherent change that, where necessary, disrupted faculty beliefs as to how well prepared their candidates were for teaching (Peck, et al., 2010; Peck et al., 2014).

Washington's teacher preparation programs also used teacher candidate performance data as a nuanced policy instrument to guide program improvement and imbue value in the edTPA work (Hill, Hansen & Stumbo, 2011). The edTPA itself was a policy instrument, used to measure the alignment of teacher candidate performance to a set of criteria that determines what beginning teachers should know and be able to do (Honig & Rainey, 2015; Hutt, Gottlieb, & Cohen, 2018; Hill et al., 2011).

The hope, in 2013, was that the edTPA and related policies would help teacher preparation programs improve their performance (Peck, et al., 2014; Peck, et al., 2010; Darling-Hammond, 2006b; Cody, 2012). Indeed, five of the top ten 2013 articles for EdWeek, a leading national education newspaper, discussed the topic of teacher preparation improvement (Sawchuk, 2013d). Ratner and Kolman (2016) found that

teacher educator attitudes toward edTPA factored into their pedagogy and the instructional supports they provided teacher candidates.

Building a culture of collaboration over time offered Washington faculty the opportunity to feel ownership and value in their practice[1] (Peck et al., 2014). Mitescu-Reagan et al. (2016) noted that five of six states they studied had field-tested the edTPA for a duration of at least four years before the 2013 implementation. Washington's field test included state and university representative consultations with SCALE, with revisions in the edTPA handbooks (Pecheone & Whittaker, 2013a; Siera, 2013) and revisions to the assessment policy itself resulting from teacher educator and organizational input (Pecheone & Whittaker, 2013b). The bottom up approach of developing a culture of collaboration leveraged faculty knowledge, and provided support with the hope that faculty would align pedagogy in a fundamental shift toward using the edTPA for program improvement (Peck et al., 2009; Peck & McDonald, 2013).

In comparison, New York's top-down policy approach allowed little room to accommodate preparation program ownership of the assessment and process and faced severe headwinds during implementation. Pushback centered around: 1) a passing standard that opponents stated was designed to fail 40% of teacher candidates, 2) the one-year state implementation timeline that many felt gave candidates insufficient time to prepare, and 3) questions of fairness of the test to teacher candidates (Sawchuk, 2014). As a result of the pushback, the New York Board of Regents delayed the

[1] In this context, the term "faculty" is used to refer to all teachers and field personnel or field supervisors associated with a teacher preparation program.

edTPA roll-out and the licensure requirement for a full year (Mitescu-Reagan, et al., 2016; Sawchuk, 2014).

The teacher preparation program actions described above indicate attempts by the programs in different states to make sense of the policy, address faculty interpretations and concerns, and derive a cohesive process to respond to the state policy. As Mitescu-Reagan et al., (2016) noted, it appears that those states that had longer pilot periods were considered more effective in adopting the state policy. These states appear to have structured opportunities to address faculty autonomy while integrating value for the process before the policy was instituted, which may have helped faculty identify the components of their program that already aligned to the edTPA standards as well as components that might be appropriate to change (Peck, Muzzo & Sexton, 2012).

2.2.4 *Alignment in Teacher Candidate Instruction*

Programs were balancing multiple implementation tensions simultaneously in the local, teacher educator arena. Teacher preparation programs focused on performance alignment to manage continuous improvement. Performance alignment is a coordinated effort to align performance to priorities (Honig, 2015; Coburn, et al., 2016). If faculty believe that the edTPA can inform their practice, they can use the edTPA as a "boundary object" (Peck, et al., 2014) to align their instructional approach to concepts measured by the edTPA. For example, in their study of PACT implementation in California (Peck & McDonald, 2013) and edTPA implementation in Washington (Peck et al., 2009), Peck and colleagues report that collaborative analysis of teacher performance assessment data encouraged instructors to re-examine their

coursework with "a more comprehensive understanding of what candidates needed" (Peck & McDonald, 2013).

The literature suggests that people make sense of their learning environments based on individual preconceptions, knowledge, skills, and attitudes about their learning, as well as feedback, and interactions within the environment in which they are learning (Bransford, Darling-Hammond, & LePage, 2005). Within teacher preparation programs, faculty, administrators, field supervisors, mentor/cooperating teachers, and teacher candidates individually and in groups co-construct understandings of policies surrounding their teaching practice to inform their actions (Honig, 2006; Spillane, Reiser & Gomez, 2006). Coburn et al. (2016) theorized that weak co-construction, or poor alignment, will lead to superficial change, for example, faculty changing what they teach, but not how they teach. Alternatively, strong alignment is thought to lead to fundamental change (Coburn, et al., 2016).

When a new term or process is introduced unilaterally, it can cause confusion to the existing understandings. For example, Washington state included in the edTPA rubric a definition of "student voice" that was contradictory to a more common definition of "student voice" used in language arts instruction. Conventionally, student voice in language arts refers to the student's writing style: the way students express their perspectives in writing, using tone and "flavor" (Education Northwest, 2015). The Washington state version of the edTPA defines student voice as the teachers' ability to elicit "student understanding of learning targets" (Stanford University, 2013). The conflicting interpretations regarding student voice may have affected Washington's

2012 Pilot edTPA scores, according to evaluation data conducted during the pilot (Peck, et al, 2012).

The low scores on the student voice section of the Washington pilot EdTPA reflected a conflict between the state's definition and the definitions of student voice used by faculty, teacher candidates, and mentor teachers. To alleviate the conflicting interpretations and align understandings to the state standard, edTPA leadership structured opportunities for programs to address misconceptions and make sense of the policy. Data from Peck, et al. (2012) indicates that edTPA leadership, AACTE, PESB, and the programs themselves developed faculty trainings as well as a discussion board and promoted conversations that allowed for faculty to develop a common understanding of, among other factors, student voice in the context of the edTPA. These communication strategies were employed as a response to the dissonance engendered by introducing terms that were contrary to pre-existing community knowledge and served as a tool to align understandings to the edTPA standards to, ultimately, program instructional performance.

Additionally, this increase in communication reflected attempts by PESB and the WACTE to attend to their constituents' feedback in order to develop common knowledge of the definition of student voice as it pertained to the edTPA. Data from Peck, et al., (2012) indicate that teacher educators were able to make mid-course corrections in response to student voice, given the inevitable vagaries and oversights that emerge in the course of an implementation.

Given these attempts by state leadership to hold teacher preparation programs accountable and to align teacher preparation curriculum and program supports with the

expectations in the edTPA, the question remains, to what extent have teacher preparation programs changed? The present study attempted to understand the outcome of the edTPA implementation, specifically the extent to which faculty perceive that the expectations for edTPA, program infrastructural supports like time and opportunities to collaboratively analyze data, teacher candidate instruction, and college missions are aligned. It is important to understand how the faculty interpreted the implementation of the edTPA at the local level, if the goal of implementation is to drive continuous improvement (Miller & Patricio, 2015; Peck, Galucci, & Sloan, 2010; Honig, 2006; Weatherly & Lipsky, 1977). By understanding the extent of alignment that occurred during Washington's edTPA's implementation, we can estimate whether the outcome led to minimal improvement, superficial change, uneven implementation, or fundamental change. Additionally, we may understand whether more alignment or more accountability would be necessary to engage and/or sustain fundamental change incurred by innovations of new tools and techniques that synthesize edTPA core concepts with current teacher educator practices.

2.3 Alignment-Accountability Matrix as Conceptual Frame

In order to understand the manner in which the edTPA's implementation in Washington influenced teacher preparation, I look to Coburn, et al. (2016) for a conceptual frame.

The present study used the alignment/accountability conceptual framework drawn from policy implementation theory to investigate the complexity within Washington State's implementation of a policy, RCW28A.410.210, calling for legislature to 1) adopt teaching standards, 2) adopt a teacher performance assessment,

and 3) state the manner in which the teacher performance assessment will be used to report teacher preparation program quality ("edTPA policy").

Coburn, et al. (2016) note two findings from implementation research, to guide the construction of their framework: 1) National and state policy can influence instructional practice, and 2) "Local implementation of state and national policy depends on alignment between those policies, district policies, and the educational infrastructure that has developed to support teachers' learning about and compliance with policy" (Coburn, et al., 2016). Their matrix was designed to consider influences between the power dynamics of Common Core State Standards accountability and standards policies and K-12 teacher practices given existent teacher learning processes. The matrix highlights the manner in which teachers might align practice to policies, given the infrastructure supporting local interpretation, and given their individual interpretations. The "Alignment" dimension refers to the manner in which learning processes are supported, to provide an alliance the policy's message. The "Accountability" dimension refers to the policy and accountability instruments used during implementation. The four quadrants represent teacher responses given the interaction of learning processes and power dynamics in play.

The alignment / accountability matrix holds promise for analyzing the manner in which faculty in Washington state teacher preparation programs responded to the edTPA policies. Consistent with the literature, we might expect faculty in higher education institutions to respond to policies and their implementations with varying amounts of change in their practice, dependent on existent knowledge and support for common interpretations provided by their environmental infrastructure (Coburn, et al.,

2016; Spillane, 2009). However, the response itself might take different forms, given how appropriate it is perceived within the environment. For example, rather than protesting with picket signs (although in a rare act, faculty in New York did picket the edTPA), faculty in higher educational settings might be expected to have scholarly debates, conduct and publish research that questions policy effectiveness, engage with state legislature in redesigning or eliminating the policy, and write books to position their stance.

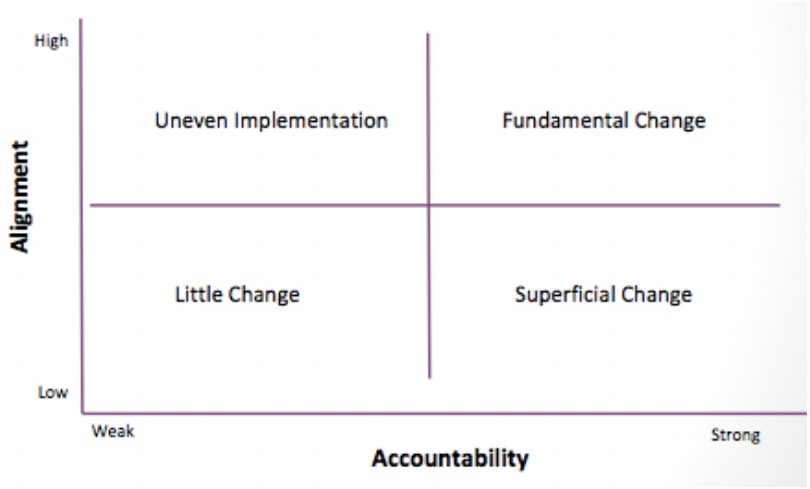
Current implementation theory calls for studies verifying that fundamental change in teaching practice occurs when there is a balance of strong accountability measures and high alignment between policy, people, and environmental infrastructure (Coburn, et al., 2016, Honig, 2015). Figure 1 (derived from Coburn et al., 2016) provides a hypothetical framework for considering how alignment and accountability interact in the context of policy implementation (Coburn et al., 2016).

The framework compares accountability instruments required, such as high stakes testing, with alignment of local practices, such as professional development, instructional practices, or administrative support (e.g., time, use of data to analyze practice), to reach the policy's goals, in order to understand the change that may have occurred. According to Coburn et al.'s (2016) model, implementation with low alignment of implementer practices to policy goals and low accountability for meeting those goals will result in minimal improvement. Teacher educators receive mixed messages and will interpret the policy in varying ways with lack of support to collectively interpret the policy's message or with pressure to change existing practices.

Faculty experiencing strong accountability measures with low alignment will be more likely to superficially change their practice, interpreting new policies within the context of preexisting knowledge with little environmental infrastructure to support fundamental change. For example, they might change instructional content, but not pedagogy (Coburn, et al., 2016). High accountability measures accompanied by low alignment have also been found to increase resistance among teachers (Coburn et al., 2016; Peck et al., 2010). For example, in K-12 educational reform research, for districts where environmental infrastructure and educational policies were aligned, but policy makers provided low accountability measures, results showed uneven implementation (Coburn, et al, 2016). There is more support but less incentive to change practice. Only those who are self-motivated and feel agency to make changes to their pedagogy will drive forward.

Finally, fundamental change is most likely for faculty where leadership, curriculum, assessment, and professional development are focused on aligning to new policies, with a keen eye on coherence among these, and where strong accountability measures are provided by leadership that faculty trust to guide them, (Lane, 1987; Coburn, et al., 2016). It is within this conceptual frame, emphasizing the interactions of alignment and accountability, that the current study sought to illuminate the outcome of the edTPA implementation. Consistent with such a view, Peck et al. (2014) reported fundamental program improvement occurred when faculty were provided opportunities to collectively analyze data that increased alignment.

Figure 1. Implementation



The present research study used the conceptual framework above, to analyze Washington's implementation of the edTPA and to analyze Washington state's policy implementation outcomes through the perspective of faculty in teacher preparation programs that implemented the policy. Specifically, the study analyzes the political dynamic of alignment in play during the Washington State implementation of the accountability policy for the edTPA, and it surveys teacher preparation program faculty involved with the edTPA about perceived changes in the program, the way that edTPA concepts were integrated into instruction, and how they see the edTPA aligning with the values of the program.

Although other theories may apply, policy implementation theory provides a theoretical framework for understanding how policy dynamics like alignment practices to policy goals or accountability mechanisms relative to the policy's objectives (as elements of a conceptual framework) interact at the local level (Weatherly & Lipsky, 1977, Lane, 1987, Honig, 2006, Coburn et al., 2016). It is necessary to understand the extent of change and factors that have affected changes that occurred in Washington's

teacher preparation programs in order to understand the policy's effectiveness to date and to specify opportunities for program improvement that exist for the future.

In what follows, I describe the methods used to apply the alignment / accountability framework within the present study. This study uses the alignment/accountability conceptual framework drawn from policy implementation theory to investigate the complexity within Washington State's implementation of a policy, RCW28A.410.210, calling for its legislature to 1) adopt teaching standards, 2) adopt a teacher performance assessment, and 3) state the manner in which the teacher performance assessment will be used to report teacher preparation program quality ("edTPA policy"). Specifically, the study analyzes the political dynamics of alignment in play during the state's implementation of the accountability policy for the edTPA, and surveys teacher preparation program faculty involved with the edTPA about perceived changes in their programs, the way that edTPA concepts were integrated into instruction, and how they see the edTPA aligning with the values of their program.

Chapter 3. RESEARCH DESIGN AND METHODOLOGY

3.1 INTRODUCTION

My research focused on three variables that influenced policy implementation results for WAC 181-78A-270 (2014) and influenced adoption and sustainability of the policy: 1) the extent to which the teacher preparation programs provided environmental infrastructure supports for faculty adoption of the edTPA, 2) the extent to which faculty (in "Foundations" and "Methods" courses) and field personnel ("Field") incorporated core concepts measured by the edTPA into their instruction, and 3) the extent to which the edTPA's core concepts were perceived as aligning to teacher preparation program values as represented in programs' mission statements.

3.2 PURPOSE

I used Coburn et al.'s (2016) conceptual frame described in the previous chapter to focus on the interaction between alignment and accountability (Coburn et al., 2016). Specifically, this study sought to identify faculty perceptions of the extent of alignment (if any) of the *core concepts* measured by the edTPA (planning, instruction, and assessment) at the teacher preparation program and instructional levels, and the extent to which faculty and field personnel believe the edTPA's core concepts are aligned to their college's mission.

3.3 Research Questions

The primary research question for this study was: What are differences between perceptions of Program Alignment, Instructional Alignment, and Mission Alignment among faculty? Secondary research questions are as follows:

- How do Methods, Field, and Foundations faculty differ in their perceptions of program-level changes in response to the edTPA??
- How do Methods, Field, and Foundations faculty differ in terms of instructional changes to address the expectations of the edTPA?
- To what extent do Methods, Field, and Foundations faculty differ in their perceptions about whether the edTPA is aligned to the teacher preparation program's mission?

The study examined whether faculty viewed program practice, faculty instruction, and program mission as aligned to the core concepts of the edTPA, planning, instruction, and assessment. Further, this study examined beliefs regarding the extent to which program administration incorporated core concepts of the edTPA (planning, instruction, assessment), as well as the edTPA standards, into infrastructural elements such as faculty development time and teacher candidate instruction. Finally, the study examined how Methods, Foundations and Field faculty viewed their program's mission in terms of alignment to the core concepts of planning, instruction, and assessment as depicted in the state standards for teacher candidate learning, and as measured by the edTPA.

Differences in the perspectives presented by faculty who teach instructional and assessment methods (Methods), faculty who teach educational foundations courses (Foundations), and faculty who support teachers in the field (Field) across institutions was expected. The rationale for expecting such differences is that the instruction of Methods and Field faculty tends to be based on application of knowledge in performance, or "clinical" situations. However, Foundations faculty may tend to focus more on comprehending research on best practices, reading research that examines, for example, education's role in society, motivation, the influence of culturally competent instruction, history of education, and social justice. The edTPA is a performance exam, and may be seen as a natural integration in Methods courses and Field work. This suggests that Methods and Field faculty may be more likely to incorporate edTPA core concepts and standards in their instruction and assessments than do Foundations faculty.

The study provided several opportunities for comparison. The study examined whether Methods and Field faculty perceived more program adjustments being made to incorporate the edTPA than did Foundations faculty. The study compared Methods, Field and Foundations faculty perceptions of instructional adjustments in their own courses to incorporate the edTPA expectations. The study compared Methods, Field and Foundations faculty response differences regarding how aligned they perceived the edTPA to be with their programs' mission, as well as how the edTPA scores have affected their programs' reputation. For example, because their work is more embedded in issues of instructional practice, Methods and Field faculty may perceive the edTPA as re-affirming their programs' priorities, whereas the Foundations faculty

may perceive the edTPA as negatively impacting their programs' priorities because it may push programs toward greater emphasis on practice overall, and less emphasis on nuanced components of teaching not measured by the edTPA. Finally, the study examined the Methods, Foundation, and Field faculty perceptions, as a group, toward alignment of program, instruction, and mission.

3.4 Research Design

In this study I used a quantitative, exploratory questionnaire research design with a targeted sample of teacher preparation program faculty and field personnel in Washington teacher prep programs. I explored perceptions of changes in faculty development efforts, instruction of teacher candidates, and program mission in order to understand the manner in which programs sustained the edTPA policy adoption. A repeated measures analysis of variance (ANOVA) was used to investigate faculty perceptions on various aspects of implementation across pilot institutions. Faculty role (that is, Methods faculty, Foundations faculty, and Field personnel roles) constituted the independent variable, and faculty responses to questions across various domains of implementation that were constructed a priori – changes in program, instruction, and mission – constituted the dependent variables in the repeated measures analysis of variance (See Table 3.1 below, "Questionnaire Scale Items").

Chronbach's alpha and exploratory factor analyses were used to investigate inter-item reliability and validity of the a priori scales (program, instruction, and mission) within the survey. Four questionnaire items were not part of the scales and were analyzed separately. Descriptive statistics (means and standard deviations) for

each item were analyzed and summarized (see Table 3.2 below, "Questionnaire Individual items").

3.5 SETTING

I surveyed teacher preparation program Methods, Foundations, and Field faculty from the seven universities that originally piloted the edTPA in 2012, in Washington State. It was important to limit the sample to the original programs that piloted the edTPA because those programs shared a common baseline in time – they implemented the pilot test during academic year 2011-2012, and thus had the same amount of time to adjust instructionally. These universities had similar amounts of time prior to the present study to prepare their program faculty to include edTPA instruction and integrate edTPA testing procedures within program processes. Additionally, individuals from all seven programs provided information to the state as they developed program supports and procedures for universities who would subsequently be implementing the edTPA.

3.5.1 *Pilot Study for the Questionnaire*

A pilot questionnaire was developed and reviewed for language, representation, and cultural sensitivity with a targeted sample of three professionals who worked in one of the universities studied. Methods faculty, Foundations faculty, and Field faculty roles were represented in the pilot. Questionnaire structure and items were revised based on pilot subjects' perceptions of question clarity and comprehensiveness. For example, there was some confusion as to whether "standards" meant the core concepts measured by the edTPA, or the passing standard set by states. As a result, a definition of terms was added to the instructions. A pilot participant suggested that a question about

whether faculty perceived the edTPA adoption as a barrier for teacher preparation programs to overcome or as an opportunity to learn would increase researchers' understanding of the effectiveness of the implementation. A second participant commented independently on the same topic as important. As a result, the question was added to the questionnaire.

3.5.2 *Primary Study*

Both the pilot and final questionnaire had 24 questions in total. Eighteen test items were ordinal, Likert-structure rating scales (1-5), with labels only on the anchors. Directions asked participants to note the extent to which they agreed with written statements, on a scale of 1-5, with one being strongly disagree and five being strongly agree. Of the eighteen questions, six questions constituted each scale: Program, Instruction, and Mission. Program questions centered on how, if at all, the teacher preparation programs arranged their structure, faculty professional development, and resources to incorporate the edTPA for program improvement purposes. The Instruction questions focused on how, if at all, individual faculty incorporated edTPA standards into course structure and activities. The Mission questions focused on how faculty perceive the criteria measured by the edTPA as aligning to the values of their institution and how the edTPA implementation has since impacted their program's reputation. A subsequent, exploratory factor analysis confirmed in detail the number of factors contributing to variance across the three scales. Questions related to these three scales are presented below in Table 3.1, Questionnaire Scale Items.

Additionally, three questions on the questionnaire were continuum-based, asking participants to mark the point on the continuum of 0 to 100 that most closely

represented their program. One question was multiple choice, asking whether the passing standard was too low, too high, or just right. The locations on the continuum and the multiple-choice question were analyzed independently.

Finally, two additional items collected demographics of the respondents, asking role, and number of years working in one or more teacher preparation programs.

3.5.3 *Sample & Data Collection*

Respondents were a targeted sample of Methods, Foundations, and Field faculty currently employed by the seven institutions that initially piloted the edTPA and responsible for instructing teacher candidates in programs that implemented the edTPA. Due to turnover beyond the control of this study, and in order to get as many responses as possible, it was not a prerequisite that respondents had to have been working from the time of the pilot. Those considered Field faculty held positions that included but were not limited to assessment coordinators, field directors and field supervisors. The target population of Methods, Foundations, and Field faculty in these seven institutions was estimated to be 300, with assumed equal distribution across the institutions. Methods and Field faculty were expected to respond similarly given that both groups' instruction was assumed to be more clinical in nature than that of Foundations faculty. Foundations faculty were expected to be less likely than Methods and Field faculty to integrate edTPA components into instruction.

Given a target population of 300, and a web-based questionnaire, the expected response rate was 24% (Bartel Sheehan, 2001), or a surveyed population of 72 people. Estimates indicated that eleven participants from each faculty group would provide 80% power to detect a two-point difference in response between groups on a scale of

one to five points, assuming the standard deviation is 1.67 points, and given the additional expectation that faculty from each of the three roles would respond in equal proportions. In actuality, forty-five participants in total responded to the survey: eleven Foundations faculty, sixteen Methods faculty, and eighteen Field faculty, with a response rate of 15%. Thus the study was considerably underpowered from the start and must be viewed as exploratory.

3.5.4 *Data Collection*

The sample population was surveyed using an online questionnaire (see Table 3.1 Questionnaire – Scale Items and Table 3.2 Questionnaire – Individual Items). A pamphlet was emailed to a member of the leadership of WACTE, asking him to forward the pamphlets to teacher preparation program directors of the seven colleges of education. I then contacted teacher preparation program directors through an introductory email with a consent form and anonymous link to the questionnaire attached. Within a two week period, only 36 responses were received. I followed up with a phone call to teacher preparation directors, explaining the study and anonymity of participants' responses, and then asked directors to forward the email again, to maintain confidentiality of potential participants.

Participants who entered the website to take the web-based questionnaire and clicked on the "Continue" button were considered volunteers. Terms signifying consent were noted both in the email and reiterated on the first page of the web-based questionnaire. Participants who clicked on the "Next" button after reading terms of consent were assumed to agree to the terms. In order to motivate responses, I included

an incentive of a lottery entry for a $50 Starbucks gift card for participation. The questionnaire had 22 survey questions (18 Likert-scale items, three continuum items, and one multiple choice item), and two demographic questions, with questions designed to minimize cognitive load.

3.6 Data Analysis Strategy

Scale evaluation.

There were six questions for each a priori scale, Program, Instruction and Mission, totaling a maximum of thirty points per scale (six 5-point Likert items in each scale). I conducted a Chronbach's alpha test to look for evidence of internal consistency of scores within each scale. A Cronbach's alpha test (see results below) was an appropriate measure of internal consistency because it assesses the strength of the correlations among items. A repeated measures ANOVA was conducted using responses across the three a priori scales. Then I used an exploratory factor analysis (EFA) as a test for evidence of internal validity, to see how the items loaded onto underlying theoretical constructs. The EFA distributed the 18 items across six constructs that differed somewhat from the three a priori scales, and an additional repeated measures ANOVA was conducted using responses on the new constructs. Questions that were independent of the scales (the final four questions) were analyzed separately by comparing mean scores.

Table 3.1 Questionnaire Scale Items

Directions: Rate the amount that you agree with each statement on a scale of 1 to 5, with 1 being strongly disagree, and 5 being strongly agree.

Definitions:
Teacher Preparation Program - the specific program within the college of education of [in?] which you primarily participate (examples: Alternative Routes, Elementary, Secondary). Participants who spend an equal amount of time with each program should consider the programs together.
Standards - In the items that follow, the term "standards" refers to the expectations from the edTPA rubrics.

Note: items seen by the faculty included a scale of one to five, with the following anchors. These statements ask about the teacher preparation program in which you teach. The next six statements refer to program level responses to adoption of the edTPA.

1. The program for which I teach has arranged its curriculum [to accommodate the edTPA].
2. Staff and faculty time is allocated so we can assist with one or more edTPA-related procedures.
3. My teacher preparation program uses the edTPA scores as one data point to internally evaluate our own performance.
4. My teacher preparation program incorporates the edTPA rubric concepts in activities to guide faculty practice.
5. My program has changed its processes to incorporate the tasks in the edTPA.
6. My program uses the edTPA scores to select professional development activities to guide faculty practice.

The next six statements refer to instructional responses to adoption of the edTPA.

7. The standards in the edTPA are relevant to the courses I teach.
8. I revised my courses to align to standards measured in the edTPA.
9. I discuss edTPA standards in my classrooms with regards to candidate performance at least once a week.
10. I reinforce edTPA standards through instructional activities at least once a week.
11. I regularly reinforce edTPA standards through homework assignments.
12. I reinforce edTPA standards through course assessments.

The next six statements refer to the relationship between your program's values/mission and the edTPA.

13. My teacher preparation program is known for producing candidates whose practice is aligned to researched best practices.
14. The demands of the edTPA are consistent with and support my program's mission.
15. My program's mission is consistent with the expectations of the edTPA.
16. My program's reputation has strengthened since the edTPA's implementation.
17. My program's mission has changed as a result of the implementation of the edTPA.
18. The edTPA has improved the reputation of my program.

Table 3.2 Questionnaire Individual Items

The next three items present continua. Slide to the area that most closely represents where your program fits on each continuum.

Range: 0 ----------50 ----------100			
1	The publication of test scores has negatively impacted my teacher preparation program.	…	The publication of edTPA scores has had a positive impact on my teacher preparation program.
2	Budgetary requirements for the edTPA implementation have negatively impacted other initiatives in my program.	…	Budgetary requirements for the edTPA implementation have had little or no impact on other initiatives in my program.
3	People in my program treated the edTPA as a barrier to overcome.	…	People in my program treated the edTPA as an opportunity to learn.
4	I think the passing standard is A. Too low B. Just right C. Too high		

Chapter 4. ANALYSIS OF THE DATA

4.1 INTRODUCTION

The purpose of this study was to investigate Methods, Foundation, and Field faculty perceptions of the relationship between teacher preparation program practices with the standards reflected in the edTPA, faculty's emphasis of the edTPA's core concepts in their instruction with teacher candidates, and faculty's perception of how well the edTPA and its standards align with preparation program mission. I explored perceptions of alignment between program support, teacher candidate instruction, and college mission with edTPA core concepts. In this chapter, I describe the data analysis and summarize the results of the study.

4.2 DATASET

The original dataset included 18 questions separated into three a priori scales: 1) Faculty and Field personnel perception of program infrastructure to incorporate the edTPA standards (see Table 4.1, "Program Items Means"); 2) Faculty and Field personnel perception of instruction of core constructs measured by the edTPA (see Table 4.2, "Instruction Items Means"); and 3) Faculty and Field personnel perception of whether their program's mission aligned to edTPA's core constructs (see Table 4.3, "Mission Items Means"). Questions 19-21 were based on a continuum and analyzed separately from the scales due to their structural difference (see Table 4.4, "Continua

Items Means"). Question 22 was a multiple-choice question and was also analyzed individually for the same reason (see Table 4.5, "Multiple Choice Item Responses").

4.3 Analysis of A Priori Scales

Forty-five participants responded to the survey: eleven Foundations faculty, sixteen Methods faculty, and eighteen Field faculty. Faculty self-identified as having 1-5 years of experience (M=3.5), 11-15 years experience (M=3.6), or 16-20 years experience (M=3.5). A single average score for each scale as defined a priori -- Program, Instruction, and Mission -- was derived for each participant, in order to maintain power (by retaining subjects even though they may have skipped an item) but still account for items not answered. Each participant's number of responses per scale was counted. A total of 6 responses were possible per scale, per participant. Participants who submitted three or more responses per scale were included in the scale's statistical analysis. Participant scores were averaged per scale, using the number of responses submitted per scale as the denominator. For example, if a participant rated item 1 as 1, left item 2 blank and responded 4, 4, 4, and 2 for questions 3-6 (Program scale), their scale score for Program was calculated as 3 (15 points divided by 5 responses).

A repeated measures analysis of variance (ANOVA) was conducted to compare the effect of faculty role on perception of preparation program's incorporation of edTPA constructs in program level policies, perception of use in instruction of constructs measured by the edTPA, and perception of program mission alignment to edTPA constructs. Means across all participants for each item are presented in Tables 4.1, 4.2, and 4.3.

Table 4.1 Program Items Means

Number	Item	Mean
1	The program for which I teach has arranged its curriculum to prepare candidates for the edTPA implementation.	3.97
2	Staff and faculty time is allocated so we can assist with one or more edTPA-related procedures	3.73
3	My teacher preparation program uses the edTPA scores as one data point to internally evaluate our own performance.	4.10
4	My teacher preparation program incorporates the edTPA rubric concepts in activities to guide faculty practice.	3.54
5	My program has changed its processes to incorporate the tasks in the edTPA.	3.90
6	My program uses the edTPA scores to select professional development activities to guide faculty practice.	2.52

Table 4.2 Instruction Items Means

Number	Item	Mean
1	The standards in the edTPA are relevant to the courses I teach.	3.94
2	I revised my courses to align to standards measured in the edTPA.	3.54
3	I discuss edTPA standards in my classrooms with regards to candidate performance at least once a week.	2.60
4	I reinforce edTPA standards through instructional activities at least once a week.	2.81
5	I regularly reinforce edTPA standards through homework assignments.	2.85
6	I reinforce edTPA standards through course assessments.	3.5

Table 4.3 Mission Items Means

Number	Item	Mean
1	My teacher preparation program is known for producing candidates whose practice is aligned to researched best practices.	4.54
2	The demands of the edTPA are consistent with and support my program's mission.	3.71
3	My program's mission is consistent with the expectations of the edTPA.	3.81
4	My program's reputation has strengthened since the edTPA's implementation.	2.75
5	My program's mission has changed as a result of the implementation of the edTPA.	2.48
6	The edTPA has improved the reputation of my program.	2.65

On items 19-21, which were scaled 0-100, faculty and field personnel responded with average scores in the mid-range of the continuum (means 59, 59, and 58, respectively) when asked if publishing edTPA scores and edTPA budgetary requirements were a negative or positive impact to their program, and whether the edTPA was a barrier or opportunity (see Table 4.4 "Continua Items Means", below).

Table 4.4 Continua Items Means

Directions: Slide to the area that most closely represents where your programs fits on each continuum.

Range: 0 ----------50 ----------100			
1	The publication of test scores has negatively impacted my teacher preparation program.	The publication of edTPA scores has had a positive impact on my teacher preparation program.	Mean: 59
2	Budgetary requirements for the edTPA implementation have negatively impacted other initiatives in my program	Budgetary requirements for the edTPA implementation have had little or no impact on other initiatives in my program.	Mean: 59
3	People in my program treated the edTPA as a barrier to overcome.	People in my program treated the edTPA as an opportunity to learn.	Mean: 58

Table 4.5 Multiple Choice Item Responses

I think the passing standard is

A.	Too low	(30% of responses, *N*=14)
B.	Just right	(66% of responses, *N*=31)
C.	Too high	(4% of responses, *N*=2)

4.4 Results Of Analysis Of A Priori Scales

Raw item data were analyzed within the three a priori scales to test each item's ability to contribute toward reliability of the questionnaire's scores and to provide evidence of validity in order to meaningfully interpret test scores.

A Cronbach's Alpha test was used to measure internal consistency, or the interrelatedness of each scale's scores: Program, Instruction, and Mission. The alpha test was used as a reliability estimate to identify the extent to which items in the questionnaire's scales measure the same construct and to identify amount of measurement error in the survey (Taylor, 2013; Tavakol & Dennick, 2011). Acceptable values of alphas range from .70 to .95 (Tavakol & Dennick, 2011). The Program subscale consisted of 6 items ($\alpha = 0.78$), the Instruction subscale consisted of 6 items ($\alpha = 0.90$), and the Mission subscale consisted of 6 items ($\alpha = 0.72$). Given such internal consistency, I proceeded with analysis of the a priori scales. Given that Mauchly's test

of sphericity[2] was not significant (p>.62) (implying that desired sphericity is likely present), the following results assume sphericity and are not adjusted.

A total of 45 participants in three roles, Foundations faculty (N=11), Methods faculty (N=16), and Field faculty (N=18), responded to the questionnaire. The independent variable, faculty role, was used to examine possible differences by role in perspective toward alignment of program facilitation of edTPA-related activities, instruction inclusion of edTPA's core concepts, and alignment of the preparation program's mission to the edTPA's core concepts.

The total mean score for the Program responses was 3.84, with a standard deviation of 0.634, n=45 (see Table 4.6 "Descriptive Statistics").

Table 4.6 Descriptive Statistics

Construct	Role	Mean	Std. Deviation	N
Program	Foundations	3.53	0.57	11
	Methods	4.00	0.54	16
	Field	3.88	0.71	18
	Total	3.84	0.63	45
Instruction	Foundations	3.45	0.85	11
	Methods	3.44	0.89	16
	Field	3.62	0.76	18
	Total	3.52	0.82	45
Mission	Foundations	3.45	0.58	11
	Methods	3.54	0.74	16
	Field	3.75	0.58	18
	Total	3.60	0.64	45

[2] Sphericity is a necessary condition for the validity of the *F* test, as it assumes that the variance of difference for each pair of factor levels is the same (Lomax, 2007). If Sphericity is violated, the Geisser-Greenhouse test must be interpreted.

Table 4.7 Mauchley's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
SCALES	.162	69.280	14	.000	.501	.564	.200

The repeated measures analysis of variance with Role as the independent variable yielded a significant main effect of Scale (i.e., overall difference in responses across scales) as follows: $F(2,84)=3.24$, $p=.044$, $\eta^2=.072$. Pairwise comparison indicated that responses about effects of edTPA implementation on the Program scale were significantly higher than those on the Instruction scale ($p<.05$) and marginally higher than those on the Mission scale ($p<.06$). The interaction, that is, the effect of both Scale and Role collectively, did not reach significance, $F(4,84)=.81$, $p>.52$. Finally, there was no significant effect of Role on responses, $F(2,42)=0.89$, $p>.41$.

4.5 Evidence for Construct Validity

Even though Cronbach's alpha tests showed reasonable internal consistency estimates, it was necessary to examine the covariation in order to understand whether unobserved variables might be overlaying the constructs (Messick, 1989). For example, "instruction" questions included questions regarding faculty integration of edTPA criteria into classroom lessons, homework, and assessment; however, the covariation in these factors might not be explained by the construct "Instruction". For example, the covariance of assessment questions might potentially be better explained if interpreted as part of a separate "assessment" construct. An exploratory factor analysis (EFA) was

conducted in order to examine the constructs being tested and identify underlying factors whose interactions may account for covariation (Messick, 1989). The EFA could reveal whether scores discriminate between different traits (discriminant validity) or converge upon one trait (convergent validity) (Messick, 1989; Allen & Yen, 1979). The additional purpose for using an EFA was to identify items that converged upon the constructs I intended to measure in the a priori scales, providing convergent evidence for validity (Messick, 1989). The factor analysis analyzed the shared variances among items, identified the number of existent underlying constructs, and identified the items that appear to measure each construct (Tabachnick & Fidell, 2007). A varimax rotation was applied to the EFA. The varimax rotation assumes orthogonality and facilitates interpretations of the manner in which factors loaded by making their correlations unambiguous. "The goal of the varimax rotation is to simplify factors by maximizing the variance of the loadings within factors, across variables" (Tabachnick & Fidell, 2007). The varimax rotation does not change the behavior of the factors, but merely amplifies their nature so that they can be interpreted.

4.6 Exploratory Factor Analysis Results

The correlations between items ranged from -.01 to .81 (see Table 4.8, "Correlation Matrix", below). The Kaiser-Meyer-Olkin measure of sampling adequacy yielded a value of 0.71, and Bartlett's test of sphericity (df, 153) was significant, $p<.01$, indicating that the assumption of sphericity was not met, and that the Geisser-Greenhouse test should be interpreted. The factor analysis yielded more than the original 3 constructs (Program, Instruction, Mission). Six constructs with Eigenvalues

greater than 1, explained 80% of the variance[3] (see Table 4.9, "Factor Analysis", below). The items loading into each construct were grouped in order to label each of the six constructs (see Table 4.10 "Rotated Component Matrix", below).

[3] I also explored models with 3, 4, and 5 factor-groupings. The items stayed relatively stable in their groupings, and supported similar theoretical constructs. I chose to use the grouping with six factors because it explained 80% of the See Appendix A.

Table 4.8 Correlation Matrix (N=48)

Correlation Matrix

		P_Curriculum	P_Time	erformance	P_Rubrics	P_Tasks	DActivities	elevance	I_Align_St	rformance	I_Activities	signments	I_Assess	_Bpractices	PA2Mission	sion2edTPA	_Reputation	ssionChange	ImprovedRep
Correlation	P_Curriculum	1.000	0.389	0.426	0.327	0.469	0.393	-0.015	0.482	0.248	0.299	0.146	0.189	0.016	0.387	0.583	0.355	0.168	0.171
	P_Time	0.389	1.000	0.148	0.297	0.376	0.426	0.164	0.254	0.172	0.147	0.082	0.327	0.381	0.119	0.297	0.190	0.230	0.347
	P_Performance	0.426	0.148	1.000	0.560	0.211	0.383	0.417	0.691	0.359	0.483	0.363	0.446	-0.129	0.596	0.507	0.376	0.220	0.140
	P_Rubrics	0.327	0.297	0.560	1.000	0.309	0.507	0.345	0.494	0.198	0.330	0.292	0.491	-0.080	0.359	0.342	0.272	0.385	0.394
	P_Tasks	0.469	0.376	0.211	0.309	1.000	0.366	0.009	0.282	0.259	0.262	0.207	0.270	0.070	0.082	0.272	0.201	0.139	0.172
	P_PDActivities	0.393	0.426	0.383	0.507	0.366	1.000	0.177	0.401	0.296	0.365	0.243	0.362	-0.037	0.291	0.163	0.311	0.354	0.417
	I_St_Relevance	-0.015	0.164	0.417	0.345	0.009	0.177	1.000	0.572	0.422	0.444	0.492	0.660	0.018	0.383	0.154	0.188	0.551	0.341
	I_Align_St	0.482	0.254	0.691	0.494	0.282	0.401	0.572	1.000	0.641	0.571	0.449	0.637	-0.079	0.485	0.412	0.374	0.435	0.268
	I_Performance	0.248	0.172	0.359	0.198	0.259	0.296	0.422	0.641	1.000	0.814	0.595	0.521	-0.237	0.332	0.263	0.369	0.576	0.417
	I_Activities	0.299	0.147	0.483	0.330	0.262	0.365	0.444	0.571	0.814	1.000	0.781	0.630	-0.272	0.372	0.290	0.183	0.506	0.232
	I_Assignments	0.146	0.082	0.363	0.292	0.207	0.243	0.492	0.449	0.595	0.781	1.000	0.571	-0.268	0.248	0.174	0.075	0.420	0.186
	I_Assess	0.189	0.327	0.446	0.491	0.270	0.362	0.660	0.637	0.521	0.630	0.571	1.000	-0.142	0.254	0.158	0.234	0.567	0.375
	M_Align_Bpractice	0.016	0.381	-0.129	-0.080	0.070	-0.037	0.018	-0.079	-0.237	-0.272	-0.268	-0.142	1.000	0.240	0.158	-0.081	-0.276	-0.092
	M_edTPA2Mission	0.387	0.119	0.596	0.359	0.082	0.291	0.383	0.485	0.332	0.372	0.248	0.254	0.240	1.000	0.721	0.372	0.173	0.193
	M_Mission2edTPA	0.583	0.297	0.507	0.342	0.272	0.163	0.154	0.412	0.263	0.290	0.174	0.158	0.158	0.721	1.000	0.331	0.088	0.124
	M_Reputation	0.355	0.190	0.376	0.272	0.201	0.311	0.188	0.374	0.369	0.183	0.075	0.234	-0.081	0.372	0.331	1.000	0.535	0.739
	M_MissionChange	0.168	0.230	0.220	0.385	0.139	0.354	0.551	0.435	0.576	0.506	0.420	0.567	-0.276	0.173	0.088	0.535	1.000	0.672
	M_ImprovedRep	0.171	0.347	0.140	0.394	0.172	0.417	0.341	0.268	0.417	0.232	0.186	0.375	-0.092	0.193	0.124	0.739	0.672	1.000

Table 4.9 Factor Analysis

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.765	37.582	37.582	6.765	37.582	37.582	3.869	21.493	21.493
2	2.295	12.748	50.329	2.295	12.748	50.329	2.696	14.979	36.472
3	1.732	9.622	59.951	1.732	9.622	59.951	2.457	13.648	50.120
4	1.343	7.461	67.412	1.343	7.461	67.412	2.040	11.335	61.454
5	1.265	7.029	74.441	1.265	7.029	74.441	1.900	10.558	72.013
6	1.001	5.561	80.001	1.001	5.561	80.001	1.438	7.989	80.001
7	.627	3.481	83.482						
8	.553	3.071	86.553						
9	.519	2.883	89.437						
10	.366	2.034	91.471						
11	.341	1.894	93.366						
12	.294	1.634	94.999						
13	.261	1.451	96.451						
14	.227	1.263	97.714						
15	.164	.911	98.626						
16	.115	.639	99.265						
17	.080	.444	99.708						
18	.053	.292	100.000						

Table 4.10 Rotated Component Matrix

	Component					
	1	2	3	4	5	6
I_Activities	.868	.217	.074	.084	.211	-.149
I_Assignments	.843	.072	-.022	.119	.087	-.116
I_Performance	.795	.210	.343	-.111	.187	-.099
I_Assess	.703	.017	.186	.481	.042	.140
I_St_Relevance	.631	.149	.210	.380	-.387	.282
I_Align_St	.559	.462	.156	.396	.132	.010
M_edTPA2Mission	.200	.854	.126	.151	-.081	.173
M_Mission2edTPA	.091	.842	.064	.024	.253	.133
P_Performance	.293	.641	.020	.541	.061	-.161
M_ImprovedRep	.154	-.010	.904	.194	.087	.092
M_Reputation	.008	.372	.844	.062	.119	-.087
M_MissionChange	.518	-.054	.700	.213	-.002	-.036
P_Rubrics	.143	.226	.180	.813	.172	-.010
P_PDActivities	.148	.060	.280	.560	.460	.047
P_Tasks	.180	.049	.036	.142	.782	.116
P_Curriculum	.055	.530	.119	.122	.668	-.029
M_Align_Bpractices	-.230	.190	-.125	-.107	-.009	.869
P_Time	.109	.008	.221	.226	.500	.655

The first factor overlapped perfectly with the a priori Instruction scale, that is, all of the a priori Instruction scale's items loaded onto this factor. Therefore, the construct was interpreted as Instructional Alignment, because the questionnaire items that loaded onto this construct involved using standards during instructional activities to align to the edTPA: 1) I reinforce edTPA standards through instructional activities at least once a week; 2) I discuss edTPA standards in my classrooms with regards to

candidate performance at least once a week; 3) I regularly reinforce edTPA standards through homework assignments; 4) I reinforce edTPA standards through course assessments; 5) I revised my courses to align to standards measured in the edTPA; 6) The standards in the edTPA are relevant to the courses I teach. The Instructional Alignment construct explained 21.5% of the variance, by far the largest of any of the factors.

The EFA indicated that the twelve items from the a priori Program and Mission scales were distributed across five underlying factors.

Factor 2 was interpreted as Mission Alignment, because its item addressed perceptions of whether the edTPA and the preparation programs' mission are aligned and supported with self-evaluation of performance: 1) My program's mission is consistent with the expectations of the edTPA; 2) The demands of the edTPA are consistent with and support my program's mission; and 3) My teacher preparation program uses the edTPA scores as one data point to internally evaluate our own performance. The Mission Alignment construct (factor) explained 15% of the variance.

Factor 3 was interpreted as Program Reputation Change, because its items addressed perceptions of whether the preparation program's reputation had changed since it began responding to the edTPA adoption: 1) The edTPA has improved the reputation of my program; 2) My program's reputation has strengthened since the edTPA's implementation; 3) My program's mission has changed as a result of the implementation of the edTPA. The Program Reputation Change construct explained 13.6% of the variance.

Factor 4 was interpreted as Performance Measures, because its items address the extent to which the edTPA performance measures (its standards and rubrics) were used to guide faculty professional development: 1) My teacher preparation program incorporates the edTPA rubric concepts in activities to guide faculty practice; 2) My program uses the edTPA scores to select professional development activities to guide faculty practice. The Performance Measures construct explained 11.3% of the variance. This factor seems reflective of the potentially positive effects that feedback and data analysis in collaborative environment from teacher performance assessments can have on program faculty (Peck et al., 2009).

Factor 5 also reflected use of feedback from the assessment, however Factor 5 was interpreted was interpreted as Program Alignment, because its items primarily focused on programs (rather than faculty) and questioned whether the preparation programs changed their processes and curricular structure to align to the core concepts measured by the edTPA: 1) My program has changed its processes to incorporate the tasks in the edTPA; 2) The program for which I teach has arranged its curriculum to prepare candidates for the edTPA implementation. The Program Alignment construct explained 10.6% of the variance.

Factor 6 was interpreted as Program Investment. This factor consisted of two items: 1) Staff and faculty time is allocated so we can assist with one or more edTPA-related procedures; and 2) My teacher preparation program is known for producing candidates whose practice is aligned to researched best practices. The first item suggests that teacher preparation programs dedicate faculty and personnel time, which signifies program commitment to the edTPA-related procedures. The second item

implies that the program reaps the benefit of a reputation for investing in instructional practices aligned to research. It could be that the coherent underlying concept that these two factors tap when loaded together is program investment in a best practice. The Program Investment construct explained 8% of the variance.

This factor analysis should be considered with caution. Although the 6-factor model explained 80% of the variance, Factors 4, 5, and 6 only contained two items, which can lead to misinterpretations. Coupled with the small number of respondents, this survey, should be considered very exploratory.

4.7 Results of Analysis of Variance of EFA Constructs

A single score for each of the six constructs, Instructional Alignment, Mission Alignment, Program Reputation Change, Performance Measures, Program Alignment, and Program Investment, was derived per participant, again with attention to maintaining power by retaining as many subjects as possible while still accounting for items not answered. Each participant's number of responses per construct was counted. Two to six responses were possible per construct, per participant, given the variable number of items varied across constructs. Given that several constructs in the EFA included only two items, participants who submitted one or more responses per construct were included in the EFA statistical analysis. Participant scores were averaged per construct, using the count of scores submitted by the individual per construct as the denominator. For example, if a participant responded with ratings 4, 4, 4, 2, 2 and 2 for questions 1-6 (Instructional Alignment scale), their construct score for Instructional Alignment was calculated as 3.0 (18 points divided by 6 responses). Construct scores were used to compare participant Role's influences on Instructional

Alignment, Mission Alignment, Program Reputation Change, Performance Measures, Program Alignment, and Program Investment using all eighteen items.

The eighteen items were redistributed within their new constructs and analyzed a second time using a repeated measures analysis of variance. The initial ANOVA provided relationships of the three a priori scales, Program, Instruction, and Mission, while the second ANOVA was based on more nuanced and empirically derived constructs identified in the EFA. For example, while the Instructional Alignment construct provided evidence that the items converged on the same construct as the initial scale, the original scale of Mission was separated into two constructs: Mission Alignment and Program Reputation Change. Interpretations of the results of the new ANOVA could indicate whether or not there is a difference between reputation and alignment, could shed light on the possibility of a link between performance and reputation (Deephouse & Carter, 2005), and might be more sensitive to the positive effects of feedback for programs and faculty (Peck et al., 2014). The second repeated measures ANOVA would illuminate more nuanced relationships, with 80% of the covariance explained, as indicated in the EFA (see table 4.9). The second ANOVA used Role as an independent variable to compare perceptions of Methods, Field and Foundations faculty across the new Constructs (dependent variables).

4.7.1 *Role as Independent Variable*

Data were analyzed using a repeated measures analysis of variance with within-subjects constructs (Instructional Alignment, Program Alignment, Program Reputation Change, Mission Alignment, Performance Measures, and Program Investment) and a between-

subject factor of Role (Methods, Foundation, Field). Means across the six constructs are shown in Table 4.11, and results of the ANOVA are presented in Tables 4.12.

Table 4.11 Means and Standard Deviations

Construct	Role	Mean	Std. Deviation	N
INSTRUCTIONAL ALIGNMENT	Foundations	3.49667	.884077	10
	Methods	3.43333	.925391	15
	Field	3.62407	.763960	18
	Total	3.52791	.834600	43
MISSION ALIGNMENT	Foundations	4.13333	.670360	10
	Methods	3.71111	1.180978	15
	Field	4.66667	2.622178	18
	Total	4.20930	1.877214	43
REPUTATION	Foundations	2.50000	1.209530	10
	Methods	2.83333	.870595	15
	Field	3.35185	.843834	18
	Total	2.97287	.987977	43
INSTRUCTIONAL PERFORMANCE MEASURES	Foundations	3.00000	.816497	10
	Methods	3.43333	.923245	15
	Field	3.27778	1.087661	18
	Total	3.26744	.965777	43
PROGRAM ALIGNMENT	Foundations	3.80000	.752773	10
	Methods	4.36667	.581460	15
	Field	4.08333	.988909	18
	Total	4.11628	.822577	43
PROGRAM BENEFITS	Foundations	3.70000	.918937	10
	Methods	4.40000	.507093	15
	Field	4.25000	.732642	18
	Total	4.17442	.747087	43

Box's M (62.58) was not significant, p=.27, indicating that the assumption of homogeneity of covariance across groups was met.

Mauchly's test indicated that the assumption of sphericity was violated ($\chi^2(14)=69.3$, $p<.05$), therefore degrees of freedom were corrected using Greenhouse-Geisser Epsilon estimate of sphericity (ε=.5). (See Table "Role Sphericity and ANOVA Results", below.) The main effect of Constructs was significant, $F(2.5,100.3)= 11.77$, $p<.05$, $\eta^2 = .23$, indicating that some constructs received higher scores than others. The main effect of Role was not significant, $F(2,40)=1.59$, $p>.05$, despite what appears to be more variation in responses across the six constructs across the three faculty roles (Methods, Foundations, and Field). There was no significant interaction between Construct and Role $F(5.0,100.3)=1.96$, $p>.31$.

Table 4.12 Sphericity and ANOVA Results

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[b]		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
CONSTRUCTS	.162	69.280	14	.000	.501	.564	.200

Tests of Within Subjects

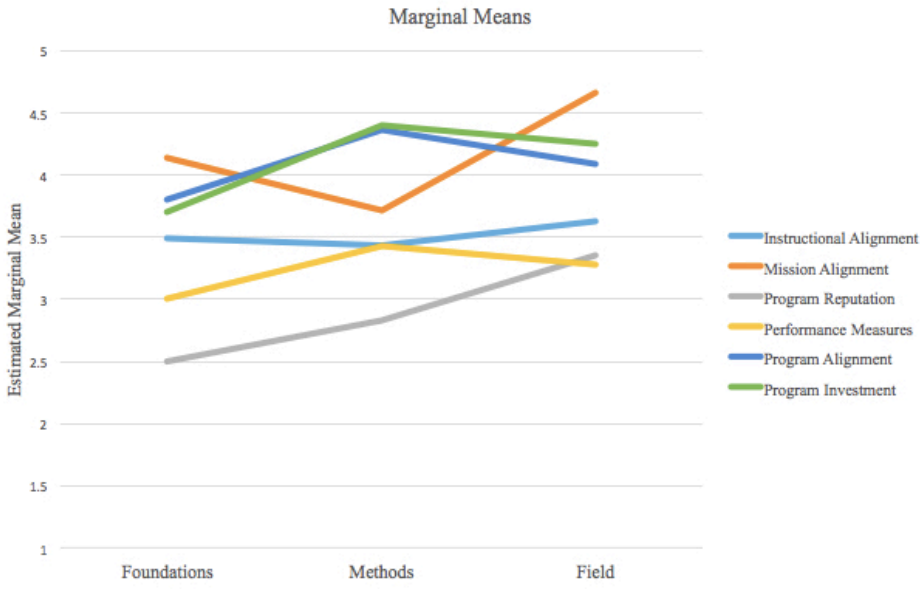
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
CONSTRUCT	Sphericity Assumed	57.988	5	11.598	11.774	.000	.227	58.872	1.000
	Greenhouse-Geisser	57.988	2.507	23.130	11.774	.000	.227	29.520	.998
	Huynh-Feldt	57.988	2.822	20.548	11.774	.000	.227	33.228	.999
	Lower-bound	57.988	1.000	57.988	11.774	.001	.227	11.774	.917
CONSTRUCT * Role	Sphericity Assumed	11.770	10	1.177	1.195	.296	.056	11.949	.613
	Greenhouse-Geisser	11.770	5.014	2.347	1.195	.317	.056	5.992	.410
	Huynh-Feldt	11.770	5.644	2.085	1.195	.315	.056	6.744	.439
	Lower-bound	11.770	2.000	5.885	1.195	.313	.056	2.390	.246
Error (CONSTRUCT)	Sphericity Assumed	196.997	200	.985					
	Greenhouse-Geisser	196.997	100.284	1.964					
	Huynh-Feldt	196.997	112.881	1.745					
	Lower-bound	196.997	40.000	4.925					

Tests of Between Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
Intercept	3273.090	1	3273.090	1405.993	.000	.972	1405.993	1.000
Role	7.407	2	3.703	1.591	.216	.074	3.182	.317
Error	93.118	40	2.328					

Follow-up pairwise t-tests showed significant differences across factors in respondent assessments of the edTPA implementation. Participant responses regarding program-related constructs, Program Alignment (M=4.12, SD=0.82) and Program Investment (M=4.17, SD=.75), did not differ from each other but both were significantly higher than responses regarding Instructional Alignment (M=3.53, SD=0.84), Program Reputation Change (M=2.97, SD=0.99), and Performance Measures (M=3.27, SD=0.97), with respective t values <.05 and d values > .24 for Instructional Alignment, d >.36 for Performance Measures, and d>.51 for Program Reputation Change. The d values represent Cohen's D, an effect size that shows the difference between standardized means (Tabachnick & Fidell, 2007). An effect size of d=0.2 is considered small, d=0.5 is considered medium, and d=0.8 is considered a large effect size (Cohen, 1992). Responses as to Mission Alignment (M=4.21, SD=1.88) were significantly higher than those regarding Performance Measures (M=3.27, SD=0.97), $t(40)=3.47$, $p<.05$, d=0.40, and Program Reputation Change (M=2.97, SD=.98), $t(40)=3.81$, $p<.05$, d=.55, but did not differ significantly from those for the other three constructs. These results are illustrated in Figure 2, separately for each faculty group.

Figure 2. Marginal Means of New Constructs by Role



4.8 Means Analysis of Individual Items

Individual questionnaire items that were not a part of the repeated measures analysis of variance were compared individually by participant role. Three such items asked participants to enter their response on a continuum of 1 – 100 (see table 4.11.1 Continuum Items, below).

Table 4.13 Continuum Items

1……50…….100

The publication of test scores has negatively impacted my teacher preparation program.	…..	The publication of edTPA scores has had a positive impact on my teacher preparation program.
Budgetary requirements for the edTPA implementation have negatively impacted other initiatives in my program.	…..	Budgetary requirements for the edTPA implementation have had little or no impact on other initiatives in my program.
People in my program treated the edTPA as a barrier to overcome.	…..	People in my program treated the edTPA as an opportunity to learn.

The three individual item scores for publication of test scores, budgetary requirements, and barrier versus opportunity were each subjected to a one-way analysis of variance with respect to faculty roles. The main effect of Role on respondent scores on the publication of test scores item was not significant, $F(2,26)=.06$, $p=.94$. Similarly, responses by Role did not differ significantly on the Budgetary requirements item ($F(2,22)=1.68$, $p=.21$) or the Barrier versus Opportunity item ($F(2,23)=1.83$, $p=.18$).

The final question was multiple choice, "I think the passing standard is…" A) Too low (score of 1); B) Just right (score of 2) or C) Too high (score of 3). A Chi Square test was performed to examine the relationship between Role and perception

about the passing standard. The relation between these variables was not significant, X^2 (2, N=47), p>.05. There were no statistically significant differences between roles in perceptions of whether the passing standard was too low, too high, or just right.

Chapter 5. DISCUSSION

Lane (1987) suggests that implementation processes do not end with the policy roll-out. Instead, implementations evolve. They evolve based on new knowledge, which in turn influences new implementations and, in the specific case of teacher performance assessments, can drive program improvement (Peck et al., 2014). In the spirit of continuous improvement, this study used Coburn et al.'s (2016) conceptual frame to examine the dimension of alignment in relation to accountability in order to understand the status of the edTPA policy implementation in Washington state as of 2017 and its perceived influence on the eight programs included in the present study (see Figure 3: Implementation Conceptual Frame Revisited).

Figure 3. Implementation Conceptual Frame Revisited

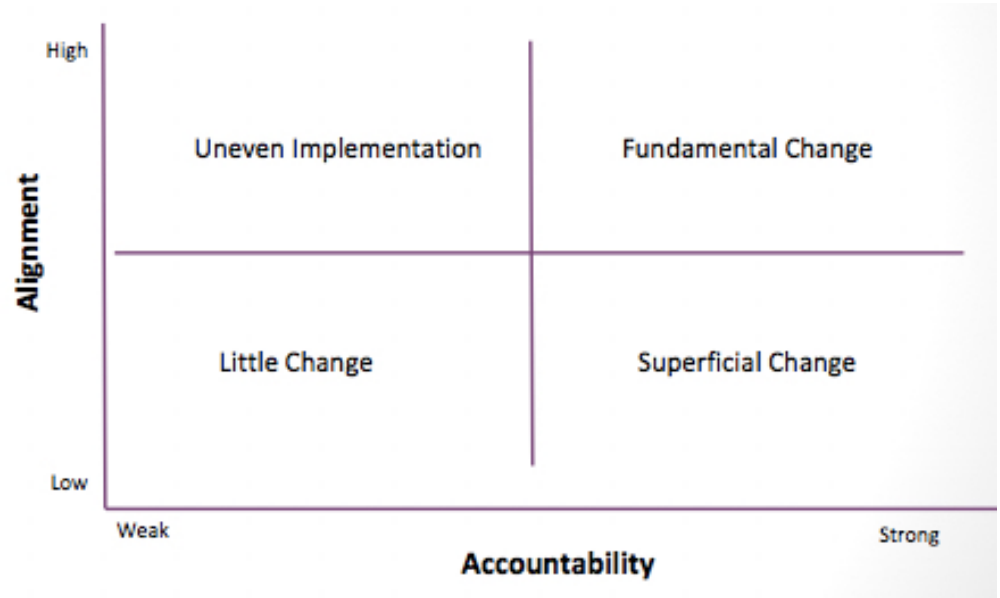


Figure 3 provides a conceptual framework, suggesting how the multiple levers of alignment and accountability (each axis in Figure 3) can influence implementation of policy and potential changes in practice. The degree of perceived program alignment with edTPA standards and the resulting instructional changes were a primary focus in the present study. Given that the edTPA is a licensure exam, accountability can be assumed to be high, leading to a prediction of at least some change at the program or instructional level (implying predicted results in the "fundamental change" and "superficial change" quadrants in Figure 3). The discussion below will focus on the six constructs resulting from the EFA of the three a priori scales, as the six constructs magnify nuances of edTPA implementation and may suggest opportunities for continuous improvement.

As stated above, the a priori factor, Instruction, overlapped perfectly with the Instructional Alignment concept that emerged from the EFA. The other five underlying factors appeared to be representative of discreet nuances within Program and Mission scales, rather than a blend of Program and Mission scales. Three of the new five factors consisted of only items from the Program scale or only items from the Mission scale - two had only a priori Program items loaded on them, and one had only Mission items loaded. The remaining two of five new factors did blend a priori Program and Mission scale items. This may suggest that a priori Program and Mission scales load into multiple discreet constructs. Moreover, the new Program Alignment construct contained items referring to observable activities, rather than the more value-laded Mission Alignment construct items. This nuanced view may be illuminating key

relationships between tactical, observable program activities and faculty attribution of alliance to program mission, which may serve as a signal of value of the edTPA policy.

Of the blended constructs, Idiosyncracies within the Program Reputation Change construct illuminated various dimensions of the a priori Mission scale by highlighting perceptions of whether the respondent's teacher preparation program's reputation has changed since responding to the edTPA adoption. The Program Investment construct tapped a blend of mission and program constructs, identifying participant perspectives of program commitment to edTPA (from the Program a priori category) and reputation for producing candidates whose practice is aligned to researched best practices (from the Mission a priori category). Similarly, Mission Alignment blended program and mission constructs to address perceptions of whether the edTPA and teacher preparation program's mission is aligned and supported with self-evaluation of performance. Both Performance Measures and Program Alignment constructs refined the Program a priori scale, reflecting the extent to which the edTPA performance measures, standards, and rubrics were used to guide faculty development activities, and whether programs changed processes or structure to align to core concepts as measured by the edTPA, respectively, as discussed by Peck and colleagues (Peck & McDonald, 2009; Peck et al., 2009; Peck et al., 2014).

5.1 How do faculty differ in perceptions of program-level changes in response to the edTPA?

Results suggest that faculty perceive that teacher preparation programs have changed processes to incorporate the edTPA tasks, and that programs changed structurally in response to the edTPA policy. Responses to the survey indicate that faculty perceive

that teacher preparation programs provided structured opportunities to interpret the edTPA core concepts and to innovate tools to integrate the concepts into courses, as reflected in responses to the following items: 1) Administration allocated time for edTPA procedures, and 2) program curriculum was restructured in order to prepare teacher candidates for the edTPA. This finding is consistent with results of a case study of one of the programs included in the present study (Peck et al., 2009), which argued that faculty had already gone through the process of innovations to fundamentally improve their instruction and align to the edTPA's core concepts. The constructs of program alignment (Factor 5) and program investment (Factor 6) had two of the three highest means among the EFA-derived factors (4.12 and 4.17, respectively). We might conclude that these two constructs reflect faculty perceptions of fundamental change (or at least substantial change) within their programs as a result of involvement with the edTPA, consistent with claims by Peck and colleagues (Peck et al., 2014).

Faculty responses to the performance measures questions (Factor 4) suggest that faculty perceive less current influence of the edTPA on their own practice, which could lead to a conclusion that faculty perceive superficial changes in how programs use the edTPA standards to provide professional development and guide faculty practice. Responses to individual items indicated faculty perceived less evidence of change in program efforts to integrate rubrics and edTPA scores into professional development activities (M=2.52 for that specific item) and less perception of change to guide faculty practice than program alignment and program investment (M=3.54). Differences between overall means for program alignment (M=4.12) and performance measures (M=3.27) were significant. Such results might indicate that teacher preparation

programs may have additional work to do to support faculty practice and develop deeper understanding among faculty of what their program's edTPA scores imply for their instruction (Peck & McDonald, 2014). Alternatively, previous research on at least one of the programs examined here suggests that considerable change in instruction occurred early in the implementation process (Peck et al., 2009), explaining perhaps why faculty perceived little current influence on their instructional practice.

5.2 How do faculty differ in terms of instructional changes to address the expectations of edTPA?

The finding of no differences across faculty roles suggests a faculty-wide coherence in perceptions regarding edTPA adoption. It may be that all faculty in the teacher preparation programs were included in program-wide opportunities to analyze teacher candidate work as researchers suggest (Peck & McDonald, 2013; Sato, 2014; Bastian et al., 2017). Perhaps faculty across the programs were provided the opportunity to make sense of the policy in the initial phases of the pilot, as found in Peck et al.'s (2009) research. Perhaps sample sizes were too small to detect differences between faculty roles. A more interesting finding was faculty perception of the nuanced differences within dimensions of teacher preparation program improvement.

The fact that faculty scored program alignment (M=4.12) and program investment (M=4.17) significantly higher than instructional alignment (Factor 4, M=3.27) is noteworthy. Faculty appear to view their own instruction as less routinely aligned to the edTPA standards when compared to their perspectives of how the program has aligned overall to incorporate edTPA standards. Faculty self-report less alignment between edTPA standards and course curriculum, classroom activities,

homework, and assessments than alignment at the program level. Faculty indicated that the edTPA standards were relevant, however they also indicated that they did not "regularly" or "weekly" reinforce the standards in assignments, homework, or class activities. There are multiple possible interpretations to this finding. The fact that faculty self-report less than weekly reinforcements of the edTPA standards in coursework may imply that faculty perceive that the edTPA standards are less relevant to their own instruction, or that faculty believe that their instruction is already in alignment with edTPA standards. Alternatively, the wording of three of the items loading on this factor asked whether faculty attend to edTPA standards "once a week" or "regularly reinforce edTPA standards," and such specificity in the time frame may have influenced faculty responses (means ranged form 2.60 to 2.85 on these individual items). Interestingly, faculty largely agreed that the edTPA standards were relevant to their courses (M=3.94), despite not reporting routine references to those standards in their instruction. However, faculty's prior work in interpreting the edTPA and innovating tools to align instruction to edTPA standards both within and between courses may have served in engaging them in the process of change prior to this study.

Faculty perceptions that edTPA standards are not regularly reinforced during instructional activities may suggest that faculty reinforce the standards, but not as "regularly" as the survey questions specified, or such a result may indicate faculty perceive alignment between teacher candidate instruction and edTPA standards already exists, and therefore the standards do not need to be foregrounded in instruction. This finding supports data from a preliminary report of the edTPA pilot implementation, which indicated that faculty did not perceive much change would occur with the edTPA

implementation, because the edTPA standards reflected the same research and best practices that they were already incorporating in their courses (Peck, et al., 2012):

> It's a necessary thing, and I think on the whole it's pretty good. I value it. But has it changed my teaching, I would not say very much. The reason is because [the edTPA] was written by what seems to be very knowledgeable people, people who read research, who are scholars in the field. And I also read that same research as I design my methods course. So most of what we teach in the methods course is congruent with a lot of the dimensions of the [edTPA]. So I've not really had to change big things in my course to align with [the edTPA] because they're already aligned for the most part.

At first glance, it would appear that improving faculty's instructional alignment with edTPA may provide opportunity for inquiry and continuous improvement within teacher preparation programs (Peck et al., 2014). As Weatherly and Lipsky note (1977), the success of a policy implementation depends on the perception of the local implementers who adopt it. However, data from previous case studies of the same universities indicate alignment prior to the current study (Peck, et al., 2012), which may have served to diminish the amount of change noted in the current study. Alternatively, it might be that: 1) the survey questions were too narrow (e.g., "once a week") for faculty to agree with honestly, or that 2) faculty believe standards are relevant but simply do not need regular reinforcement in class activities, assignments, or homework, or 3) perhaps faculty were already incorporating the standards.

Perhaps one avenue for continuous improvement would be to continue to address autonomy and local control, engage faculty in interpreting their work in light of current edTPA results, and continue to engage faculty in structured, ongoing professional development. The importance of respecting the concerns of local participants and engaging them in policy implementation was a key finding in the

development of the Performance Assessment for California Teachers, a precursor to the edTPA (Peck et al., 2010). Researchers have found that policies engendering high accountability practices like the edTPA must be met with equally strongly emphasized activities to invest in the value of local participants, and engage them during implementation, in order to infuse value, and assist in interpreting policy requirements within current, local contexts (Wenger,1998; Peck et al., 2010; Peck et al., 2009; Plecki, Elfers, & Nakamura, 2012). Peck et al. (2010) noted three key practices that assisted faculty in seeing the teacher performance assessment as an opportunity for fundamental improvement: 1) respecting and making concerns of local practitioners visible; 2) instituting regular data analysis and discussions between teacher preparation program leadership and faculty members; and 3) identifying and supporting collective and individual learning opportunities for faculty.

Faculty perception that edTPA standards are not regularly reinforced during instructional activities may indicate that an opportunity exists to revisit their concerns, provide opportunities for faculty to collaboratively analyze edTPA data, and support their collective and individual learning opportunities around their instructional practice. Multiple explanations exist to account for faculty ratings on the instruction items, but the present data cannot conclusively identify which is correct. Teacher preparation program faculty participating in the current study may have received conflicting messages when implementing the edTPA, (Coburn, et al., 2016) and may not have routinely integrated edTPA standards into their weekly instructional practice (Peck, et al., 2010), or possibly not have seen the need to do so as they were already aligned. Alternatively, faculty may be reinforcing the standards, but not as "regularly" as the

survey questions specified. If faculty ratings do reflect inconsistent messaging, increased faculty development opportunities supporting alignment of edTPA standards within routine instruction may assist in moving performance from mere compliance, or superficial instruction, to a fundamental change of continuous improvement through inquiry of practice (Peck, et al., 2010).

5.3 How does the edTPA construct align to the mission of the teacher preparation programs?

The mission alignment construct (Factor 2, M=4.21) was scored by faculty significantly higher than program reputation change (Factor 3, M=2.97) and performance measures (Factor 4, M=3.27). The mission alignment construct addressed perceptions of whether the edTPA and preparation programs' mission were aligned, and whether programs used edTPA scores to internally evaluate faculty performance. These items were scored significantly higher than performance measures, which addressed the extent to which the edTPA performance measures, its standards and rubrics, were currently used to guide faculty development activities. This may indicate that faculty perceive that their instructional practices were already aligned with edTPA standards and they do not need further professional development.

Faculty attributed limited change in their program's reputation as a result of the edTPA (responses to Factor 3). This finding is especially interesting given their strong belief that their programs are "known for producing candidates whose practice is aligned to researched best practices" (M=4.54 for that specific item in Factor 6). It could be that change in program reputation occurred, but faculty do not attribute it to the edTPA implementation. Alternatively, faculty may believe that their programs have

always been aligned to the edTPA, and that no change was necessary. It is suggested that further studies investigate the relationship between program reputation and the edTPA.

5.4 LIMITATIONS

In order to reduce measurement error, a pilot survey of the three original scales was developed and reviewed for language, representation, and cultural sensitivity with a convenience sample of three people who work in the University of Washington's teacher preparation program – one methods faculty member, one foundations faculty, and one field personnel member. The purpose of the pilot survey was to identify item comprehension issues, the range of answers I might receive, and the extent of measurement error with my questions given the manner in which the three a priori scales were originally worded. Survey structure and items were revised accordingly, with two items added in response to the pilot. Despite the pilot study, the EFA suggested the questions reflected as many as six underlying factors. It must be noted however, that the EFA discussed here should be considered very exploratory and thus interpreted with considerable caution due to the limited sample size and undesireable overdetermination of factors, that is, a small number of items per factor, which may have contributed to sampling error (MacCallum, Widaman, Zhang, & Hong, 1999). Alternative EFAs with fewer constructs should be considered. The survey itself could be improved through repeated pilots, and by adding more items that examine each of the six constructs. Although this study sought to use Coburn et al.'s (2016) conceptual framework to consider the link between accountability and alignment, this study should not be considered a test of the framework, itself.

A second limitation is the focus of the study on only the eight university programs involved in the edTPA pilot. Online universities that admit Washington state teacher education candidates were not included in the study. The online universities from out of state represent a different population because they were not held to the same accountability standards of having their teacher candidates complete the edTPA as were in-state institutions of higher education at the time of this study. Further research is necessary to identify the impact, if any, of edTPA results on online university teacher preparation. Additionally, universities that were not a part of the initial eight pilot institutions were not included in this study, in order to maintain a manageable project scope for a single investigator. However, the sample pilot institutions were intentionally targeted because of their experience in developing and piloting the edTPA. Given their involvement in the edTPA pilot, it was reasonable to assume that these universities would have had the greatest opportunity to align their program's processes and instructional activities to the edTPA and state standards for beginning teachers. This strategy may have led to response bias, because those who chose not to respond may hold different values regarding the edTPA than those who responded. To attempt to combat this, the email invited participants by stating the purpose as determining "how the edTPA has influenced the nature of practice of teacher preparation programs in Washington State." The invitation noted that participants would be free to not answer any questions, and would contain no links between email address and survey responses, in order to protect respondent identity. Finally, the invitation reiterated the need to inform policy makers and practitioners concerned with the quality of teacher education, in the hopes of drawing responses

from practitioners with multiple perspectives. Still, the question remains as to possible differences between those who chose to respond to the survey and those who did not. Future research should do more to welcome all perspectives and emphasize the importance of multiple perspectives as being part of the fabric of higher educational institutions. Additionally, future studies should invite all universities in Washington to participate, as an increased sample size could reflect significant differences in scores, and there may be additional differences between pilot institutions and those not participating in early efforts.

In addition, future research might link study participants to their university programs, which would be helpful in understanding the generalizability of the study results. There is a need for more research regarding perspectives of the entire population of teacher preparation program faculty in Washington state, as the original pilot institutions' responses may not be representative of the entire population.

Finally, the small sample size, which contributed to the exploratory nature of the factor analysis, and potentially to response bias, greatly impacted this study. The low response rate among program faculty may have been due to time, priorities, or cognitive burden. In order to motivate responses and reduce nonresponse bias, the incentive of a lottery ticket for a $50 Starbucks gift card for participation was included in the introductory email. The survey included 22 questions, which took approximately 15 minutes to complete. This short amount of time was noted on the introductory email to potential participants, to motivate engagement. Despite such efforts, the response rate was lower than anticipated. The initial power analysis suggested that a two-point difference between roles would be needed to detect a significant difference in responses

between groups. However, in the actual data, the average difference between Field (M=3.62) scores on alignment and Foundation (M=3.45) scores was 0.17, with a standard deviation of 0.82, indicating the need for a sample size of 294 participants to derive 80% power. (A smaller difference across roles requires a larger sample size to detect significance, in this case, almost the entire population of faculty in these roles in Washington). The study was thus underpowered to draw a conclusion about a difference of 0.17. Still, given similarities in overall means across groups, it is doubtful, at the time of the study, a meaningful difference in perceptions between Methods, Foundations, and Field faculty existed. Additionally, differences across Methods, Field and Foundations faculty may not have been observable because there was minimal recent reformulation of instructional practices in response to the edTPA. Simply put, if few meaningful changes in current faculty instruction were perceived to result from implementation of the edTPA, no differences in reported level of change would be observed between more practice-focused faculty (Methods and Field) and Foundations faculty. Whether this was a result of changes early in the implementation process (Peck et al., 2009), prior to the present study, is a question that cannot be answered definitively here. Future work should invite all teacher preparation programs in Washington state to participate, thus increasing the potential sample size, or utilize other ways to assess whether and which changes were made.

The small sample size also necessitated that the factor analysis be considered quite exploratory. MacCallum et al., (1999) note the importance of sample sizes larger than sixty as one factor to decrease the variability in factor loadings across repeated samples. I decided to conduct an EFA anyway, in order to see how the data loaded, and whether

or not data loaded onto unanticipated theoretical constructs. Subsequently, the data loaded onto six factors with 80% of the variability explained. When forced into five factors, four factors, and three factors, 74%, 67%, and 60% of the variability was explained respectively. I decided to use the EFA with six factors because that model explained the most variability. A six-factor model may have over-parsed the data, and contributed to weak over-determination of factors. Sampling error may have been increased because there were not enough items per construct to make the construct robust (MacCallum et al., 1999). Interestingly, three factors showed some stability, and loaded exactly the same within the four, five, and six factor models (See Appendix A). The items within the a priori construct, "Program" re-distributed as the number of factors grew, and two items cross-loaded onto multiple constructs, making interpretation difficult. Time allocated for edTPA work, and performance items cross-loaded onto other factors. Further analyses should interpret a model with four or possibly five factors. It is recommended that a broader population be surveyed in order to draw substantial conclusions.

5.5 Reflections

This study set out to describe implementation of teacher education reform that occurred at the local level, in Washington state teacher preparation programs following passage of the mandate for a teacher performance assessment and, consequently, the adoption of the edTPA. Specifically, this study sought to understand faculty perceptions of the extent of program and instructional improvements made to align to edTPA core concepts, as well as how well faculty felt the edTPA core concepts aligned to their program's mission. Coburn et al.'s (2016) conceptual framework was useful to

understand how perceived changes may have manifested in program improvement efforts. Findings support prior research supporting the ongoing use of data and structured opportunities to engage faculty in making sense of policies (Peck et al., 2009; Peck et al., 2010; Peck et al., 2013; Peck et al., 2014), as well as the importance of aligning local activities to interpretations of the policy's message within local contexts to support fundamental change (Coburn et al, 2016).

The *manner* in which teacher preparation programs structure opportunities for faculty learning counts. If the expectations of the edTPA are truly best practices for teaching and should be included in teacher preparation programs, the results of this study suggest that programs would benefit by continuing to address performance alignment in instructional activities, homework, assessments, and curriculum, which in turn stresses relevance of the standards to desired practice. Collectively deconstructing and adapting edTPA policies' messages to the local context allows faculty and program leadership the space to analyze what they are doing well, value their work, and to innovate improvements (Peck et al., 2010; Darling-Hammond, 1990). Local leadership plays a key role in navigating faculty toward autonomy over and value of the instruction they provide teacher candidates (Mitescu-Reagan et al., 2016), while simultaneously addressing political tensions of colonization of curriculum (Mitescu-Reagan et al., 2016; Evans, 2015), marginalization during high stakes accountability mandates (Evans, 2015; Coburn, et al, 2016) and a potential, perceived loss of local control (Mitescu-Reagan, et al, 2016).

This research connects to the larger issues of educational policy and policy implementation by analyzing the perspectives of the local practitioners responsible for

implementing one influential policy, the edTPA, and attempting to tap constructs that may prove resistant to fundamental change (Spillane, 2009). Policy analysts cannot assume that an implementation has ended, just as we cannot assume practitioners' enactments of a policy years after initial implementation will resemble their initial response. Moreover, policy analysts have a duty to represent multiple perspectives involved in an implementation over time, when considering its success.

Valuing practice is both an individual and a collective activity. When faculty reflect on the data individually, interpret its meanings, and collectively innovate strategies for program improvement (Peck, et al., 2009), they essentially decide which innovations are of the most value for their teacher preparation program improvement efforts. As Evans (2015) notes, surely there can be a balance of measured curricula with the autonomy to instruct other dimensions of teaching, such as motivation or social justice, as both are equally important aspects of teaching.

Teaching is an evolving process and requires faculty to consistently analyze teacher candidate work to understand what candidates are learning versus what faculty intended them to learn (Mitescu-Reagan, 2016; Darling-Hammond & Bransford, 2005). Programs could benefit by incorporating the edTPA's standards and scores into their faculty development activities to deliberately analyze their own practice and innovate improvements on a regular basis (Peck et al., 2014). Ongoing, structured opportunities to use qualitative and disaggregated quantitative data at the local level can assist in improving faculty's assessment literacy (Bastian, et al., 2017), and critical analysis of current student work (Peck et al., 2010), to lead toward fundamental shifts in teacher preparation program improvement.

REFERENCES

Allen, M. J. & Yen, W. M. (1979). *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.

Bartel Sheehan, K. (2001). Email survey response rates: A review. *Journal of Computer-Mediated Communication*, 6(2). JCMC621, https://doi.org/10.1111/j.1083-6101.2001.tb00117.x

Bastian, K. C., Lys, D., & Pan, Y. (2017). A Framework for Improvement: Analyzing Performance Assessment Scores for Evidence-Based Teacher Preparation Program Reforms. *Education Policy Initiative at Carolina*, 31.

Bransford, J., Darling-Hammond, L., and LePage, P. (2005). Introduction. In L. Darling-Hammond and J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 1-39). San Francisco, CA: Jossey-Bass.

Chung, R. R. (2008). Beyond Assessment: Performance assessments in teacher education. *Teacher Education Quarterly*, *35*, 7–28.

Coburn, C., Hill, H., and Spillane, J. (2016). Alignment and accountability in policy implementation: The Common Core State Standards and implementation research. *Educational Researcher*, 45(4), 243-251.

Cochran-Smith, M., & Villegas, A. M. (2015). Framing Teacher Preparation Research: An Overview of the Field, Part 1. *Journal of Teacher Education*, *66*(1), 7–20. http://doi.org/10.1177/0022487114549072.

Cody, A. (2012, December 14). Guest post by Ann Schulte. Ann Schulte: Teacher Performance Assessment Isn't the Answer. *Education Week.* Retrieved from http://blogs.edweek.org/teachers/living-in-dialogue/2012/12/ann_schulte_teacher_performanc.html?qs=edtpa.

Cohen J. (1992). A Power Primer. *Psychological Bulletin*, *112*(1). http://doi.org/10.1038/141613a0.

Courrege-Casey, D. (2013, October 14). Move to Performance Testing for Teachers Found Slow to Spread to Rural States. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/rural_education/2013/10/move_to_performance-based_tests_for_prospective_teachers_slow_in_rural_states.html?qs=edtpa.

Cronbach, L. J., & Shavelson, R. J. (2004). My Current Thoughts on Coefficient Alpha and Successor Procedures. *Educational and Psychological Measurement*, *64*(3), 391–418. http://doi.org/10.1177/0013164404266386

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–298.

Darling-Hammond, L. (1990). Instructional Policy Into Practice: "The Power of the Bottom Over the Top." *Educational Evaluation and Policy Analysis*, *12*(3), 339–347. http://doi.org/10.3102/01623737012003339

Darling-Hammond, L., & Bransford, J. (2005). Preparing teachers for a changing world. San Francisco, CA: Josse Bass.

Darling-Hammond, L., & McLaughlin, M. W. (2011). Policies that support professional development in an era of reform. *Phi Delta Kappan*, *92*(6), 81–92. http://doi.org/10.1177/003172171109200622

Darling-Hammond, L., & McLaughlin, M. W. (1999). Investing in teaching as a learning profession: Policy, problems and prospects. In L. Darling-Hammond & G. Sykes (Eds.), Teaching as the learning profession: Handbook of policy and practice (pp. 376-411). San Francisco, CA: Jossey-Bass.

Darling-Hammond, L. (2006a). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education* 57(2), 120-138.

Darling-Hammond, L. (2006b). Constructing 21st century teacher education; *Journal of Teacher Education*, 57(3), 300-314.

Durand, F., Lawson, H., Wilcox, K., Schiller, K. (2016). The role of district office leaders in the adoption and implementation of the Common Core State Standards in elementary schools. *Educational Administration Quarterly,* 52(1) 45-74.

Education Northwest, 2015. *"6+1 Trait Writing"*. Retrieved from http://educationnorthwest.org/traits/trait-definitions.

Evans, C. (2015). Examining Policies & Reform Agendas in Teacher Education : Historical Analysis of the Discourses & Arguments Surrounding Performance Assessments. In *New England Educational Research Organization* (p. 38). Portsmouth, NH: ResearchGate.

Hargreaves, A. & Fullan, M. (2012). Professional capital: Transforming teaching in every school. New York: NY, Teachers College Press.

Heltin, L. (2013, November 11). Setting the Bar for Teacher Licensing. *Education Week.* Retrieved from http://blogs.edweek.org/teachers/teaching_now/2013/11/setting_the_bar_for_teacher_licensing.html?qs=edtpa.

Henry, G., Campbell, S., Thompson, C., Patriarca, L., Luterbach, K., Lys, D., & Covington, V., (2013). The predictive validity measures of teacher candidate programs and performance: toward an evidence based approach to teacher preparation. *Journal of Teacher Education*, 64(5), 439-453.

Hill, B., Hansen, D., & Stumbo, C. (2011). Policy considerations for states participating in the Teacher Performance Assessment consortium. *Council of Chief State Schools Officers.* Retrieved from http://westwinded.com/lib/mats/TPACPolicyPaperwithProfiles_ForDistribution.pdf

Honig, M. (2006). *New directions in education policy implementation: Confronting complexity.* (M.I. Honig, Ed.). Albany, NY: State University of New York Press.

Honig, M.. & Rainey, L. R. (October, 2015). How school districts can support deeper learning: The need for performance alignment. Deeper Learning Research Series. *Jobs for the Future*. Boston, MA. www.jff.org. ERIC Number: ED560756. 26pp.

Hutt, E. L., Gottlieb, J., & Cohen, J. J. (2018). Diffusion in a vacuum: edTPA, legitimacy, and the rhetoric of teacher professionalization. *Teaching and Teacher Education*, *69*, 52–61. http://doi.org/10.1016/j.tate.2017.09.014

Kohoutek, J. (2013). Three Decades of Implementation Research in Higher Education: Limitations and Prospects of Theory Development. *Higher Education Quarterly*, *67*(1), 56–79. http://doi.org/10.1111/j.1468-2273.2012.00531.x

Lane, J. (1987). Implementation, Accountability, and Trust. *European Journal of Political Research*, 15(5), 527-546.

Lewis, W. D. and Young, T.V. (2013). The politics of accountability: Teacher education policy. *Educational Policy* (27), 190-216. DOI: DOI: 10.1177/0895904812472725. Retrieved from http://www.gare.cree-inter.net/sites/default/files/Educational%20Policy-2013-Lewis-The%20Politics%20of%20Accountability%20%20Teacher%20Education%20Policy.pdf

Lomax, R. (2007). *Statistical Concepts: A Second Course*. Mahwah: NJ. Lawrence Erlbaum Associates, Inc.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample Size in Factor Analysis. *Psychological Methods*, *4*(1), 84–99.

Messick, S. (1989). *Validity*. In R.L. Linn (Ed.), Educational measurement (3rd Ed.). New York: American Council on Education/Macmillan. Messick,

Meyer, H.D. & Rowan, B. (2006). *The New Institutionalism in Education*. NY: State University of New York Press.

Miller, G.R. and Patrizio, K.M. (2015). Agency as Place in Teacher Education. *The New Educator* 11, 309-321. DOI: 10.1080/1547688X.2015.1087739.

Mintrop, H. & Sunderman, G. (2009). Predictable failure of federal sanctions-driven accountability for school improvement – And why we may retain it anyway. *Educational Researcher,* 38(5), 353-364.

Mitescu-Reagan, E. M., Schram, T., McCurdy, K., Chang, T.-H., & Evans, C. M. (2016). Politics of Policy : Assessing the Implementation, Impact, and Evolution of the Performance Assessment for California Teachers (PACT) and edTPA. *Education Policy Analysis Archives*, *24*(9), 1–24.

No Child Left Behind Act of 2001, 20 U.S.C. § 9101 C.F.R. (2001).

Organization for Economic Co-operation and Development (OECD). (2005). *Teachers matter: Attracting, developing, and retaining effective teachers*. Paris, France: Author.

Paufler, N., & Amrein-Beardsley, A. (2016). Preparing teachers for educational renewal within current contexts of accountability: Reflecting upon John Goodlad's twenty postulates. *Journal of Teacher Education*. 67(4), 251-262.

Pecheone, R. L., & Chung, R. R. (2006). Evidence in teacher education: The Performance Assessment for California Teachers (PACT). *Journal of Teacher Education*, *57*(1), 22–36. http://doi.org/10.1177/0022487105284045.

Pecheone, R. L., & Whittaker, A. (2013a). Response to edTPA retake policy: dated 12/3/13. Retrieved from http://assessment.pesb.wa.gov/assessments/edtpa/wacte-documents.

Pecheone, R. L., & Whittaker, A. (2013b). Response to WACTE feedback on the edTPA: dated 8/12/13. Retrieved from http://assessment.pesb.wa.gov/assessments/edtpa/wacte- documents.

Peck, C. A., Gallucci, C., & Sloan, T. (2010). Negotiating Implementation of High-Stakes Performance Assessment Policies in Teacher Education : From Compliance to Inquiry. *Journal of Teacher Education*, *61*(5), 451–463. doi:10.1177/0022487109354520.

Peck, C. A., Singer-Gabella, M., Sloan, T., & Lin, S. (2014). Driving Blind: Why We Need Standardized Performance Assessment In Teacher Education Charles A. Peck, *8*(1), 8–30. http://doi.org/10.3776/joci.2014.v8n1p8-30.

Peck, C. A., Gallucci, C., Sloan, T., & Lippincott, A. (2009). Organizational learning and program renewal in teacher education: A socio-cultural theory of learning, innovation and change. *Educational Research Review*, *4*(1), 16–25. http://doi.org/10.1016/j.edurev.2008.06.001.

Peck, C. A., & McDonald, M. (2013). Creating "Cultures of Evidence" in Teacher Education: Context, Policy, and Practice in Three High-Data-Use Programs. *The New Educator*, *9*(1), 12–28. http://doi.org/10.1080/1547688X.2013.751312.

Peck, C. A., & McDonald, M. A. (2014). What is a culture of evidence? How do you get one? and...Should you want one? *Teachers College Record*, *116*(3), 1–27. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2014-16783-006&site=ehost-live

Peck, C. A., Gallucci, C., & Sloan, T. (2010). Negotiating implementation of high-stakes performance assessment policies in teacher education: From compliance to inquiry. *Journal of Teacher Education*, 61(5), 451-463.

Peck, C., Muzzo, M., & Sexton, P. (2012). Program implementation of the edTPA in Washington State. *Unpublished manuscript. Seattle, WA: University of Washington*.

Plecki, M., Elfers, A., & Nakamura, Y. (2012). Using evidence for teacher education program improvement and accountability: An illustrative case of the role of value added Measures. *Journal of Teacher Education*. 63(5), 318-334.

Ratner, A. & Kolman, J., (2016). Breakers, benders, and obeyers: Inquiring into teacher educators mediation of edTPA. *Education Policy Analysis Archives*. 24(35).

RCW 28A.410.270, (2013). Washington State Legislature. *RCW 28A.410.270 Washington professional educator standards board – performance standards – certification levels – teacher effectiveness evaluations – update – proposal – recommendation – requirements for professional certification and residency teaching certificate.* Retrieved from http://apps.leg.wa.gov/RCW/default.aspx?cite=28A.410.270.

RCW28A.410.210, (2013). Washington State Legislature. *RCW28A.410.210: Washington professional educators standards board – purpose, powers and duties*. Retrieved from http://apps.leg.wa.gov/RCW/default.aspx?cite=28A.410.210.

Robb, L. (2013). New angles on differentiating reading instruction: Five best practices that deserve a new chapter in the Common Core era. *The NERA Journal*, *49*(1), 13–22.

Sabatier, P., (1986). Top-down and bottom-up approaches to implementation research: A critical analysis and suggested synthesis. *Journal of Public Policy*, 6(1) 21-48. Retrieved from http://www.jstor.org/stable/3998354.

Saltmarsh, D. (2012). Voices in education: Politics and teacher education. *The Teacher Educator*, 47(3), 169-174. Taylor & Francis Group. DOI: 10.1080/08878730.2012.686404.

Sato, M., (2014). What is the underlying conception of teaching of the edTPA? *Journal of Teacher Education*. 65(5) 421-434.

Sawchuk, S. (2012, September 20). Democrats introduce bill to overhaul teacher training. *Education Week.* Retrieved from http://blogs.edweek.org/edweek/teacherbeat/2012/09/_have_introduced_companion_bil.html?qs=edtpa.

Sawchuk, S. (2013a, December 4). Performance Key on Teacher Tests. *Education Week,* 33(13), pp 1-22. Retrieved from http://www.edweek.org/ew/articles/2013/12/04/13assess_ep.h33.html?qs=edtpa.

Sawchuk, S. (2013b, December 2). Washington, New York Set Passing Bars on New Teacher Test. *Education Week.* Retrieved from http://blogs.edweek.org/edweek/teacherbeat/2013/12/not_long_ago_i_posed.html?qs=edtpa.

Sawchuk, S. (2013c, November 8). A New Teacher-Licensing Test: Where Will States Set the Bar? *Education Week.* Retrieved from http://blogs.edweek.org/edweek/teacherbeat/2013/11/group_recommends_score_on_mult.html?qs=edtpa.

Sawchuk, S. (2013d, December 19). Teacher Beat's Top 10 for 2013. *Education Week.* Retrieved from http://blogs.edweek.org/edweek/teacherbeat/2013/12/10_this_one_gets_my.html?qs=edtpa.

Sawchuk, S. (2014, April 29. Under Legislative Pressure, N.Y. Delays Licensing-Test Deadline. *Education Week.* Retrieved from *http://blogs.edweek.org/edweek/teacherbeat/2014/04/NY_delays_licensing_test_deadline.html?qs=edtpa.*

Siera, S. (2013, November 14). Letter to SCALE regarding retake policy. Retrieved from http://assessment.pesb.wa.gov/assessments/edtpa/wacte-documents.

Spillane, J Spillane, J. P. (2009). Implementation Reconsidered. In *Standards Deviation: How Schools Misunderstand Education Policy* (pp. 168–187). Cambridge, MA: Harvard University Press. Retrieved from https://ebookcentral-proquest-com.offcampus.lib.washington.edu/lib/washington/reader.action?docID=3300672&query=

Spillane, J., Reiser, B., and Gomez, L. (2006). In Honig, M.I. (Ed.) *New directions in education policy implementation: Confronting complexity*. Albany, NY: State University of New York Press.

Stanford Center for Assessment, Learning and Equity (SCALE). (2013). *edTPA*. Retrieved from https://scale.stanford.edu/teaching/edtpa.

Stanford Center for Assessment, Learning and Equity (SCALE). (2014). *By the numbers: 2014 edTPA administrative report*. Retrieved from:

https://scale.stanford.edu/sites/default/files/edTPA By the Numbers 929 Final.pdf.

Stanford University, September 2011. Teacher Performance Assessment Consortium WA TPA secondary English language arts assessment handbook: Field test pre-release handbook. Stanford: Stanford, CA.

Tabachnick, B. and Fidell, L. (2007). *Using multivariate statistics.* 5th ed. Boston, MA: Pearson Education, Inc.

Tatto, M. T. (2006). Education reform and the global regulation of teachers' education, development and work: A cross-cultural analysis. *International Journal of Educational Research*, *45*(4–5), 231–241. http://doi.org/10.1016/j.ijer.2007.02.003.

Tavakol M, Dennick R. Making sense of Cronbach's alpha. *International Journal of Medical Education*. 2011;2:53-55. doi:10.5116/ijme.4dfb.8dfd.

Taylor, C. (2013). *Validity and Validation*. New York, NY: Oxford University Press.

Weible, C. M., Heikkila, T., deLeon, P., & Sabatier, P. A. (2012). Understanding and influencing the policy process. *Policy Sciences*, *45*(1), 1–21. http://doi.org/10.1007/s11077-011-9143-5

Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge, UK: Cambridge University Press.

Weatherley, R. and Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Education Review*, 47(2), 171-197.

Wiseman, D. L. (2012). The Intersection of Policy, Reform, and Teacher Education. *Journal of Teacher Education*, *63*(2), 87–91. http://doi.org/10.1177/002248711142.

Chapter 6. APPENDIX

6.1 Factor Analysis Comparison

I explored models with 3, 4, and 5 factor-groupings. The items stayed relatively stable within their groupings, and supported similar theoretical constructs. Although I chose to use the model with six factors because it explained 80% of the variance, a 4 or 5-factor model may have been more appropriate to use to reduce measurement error. EFA results for factor groupings are below.

Table 6.1 Factor Analysis Comparison – Rotated Component Matrices

Rotated Component Matrix[a]

	Component					
	1	2	3	4	5	6
I_Activities	.868	.217	.074	.084	.211	-.149
I_Assignments	.843	.072	-.022	.119	.087	-.116
I_Performance	.795	.210	.343	-.111	.187	-.099
I_Assess	.703	.017	.186	.481	.042	.140
I_St_Relevance	.631	.149	.210	.380	-.387	.282
I_Align_St	.559	.462	.156	.396	.132	.010
A_edTPA2Mission	.200	.854	.126	.151	-.081	.173
A_Mission2edTPA	.091	.842	.064	.024	.253	.133
P_Performance	.293	.641	.020	.541	.061	-.161
A_ImprovedRep	.154	-.010	.904	.194	.087	.092
A_Reputation	.008	.372	.844	.062	.119	-.087
A_MissionChange	.518	-.054	.700	.213	-.002	-.036
P_Rubrics	.143	.226	.180	.813	.172	-.010
P_PDActivities	.148	.060	.280	.560	.460	.047
P_Tasks	.180	.049	.036	.142	.782	.116
P_Curriculum	.055	.530	.119	.122	.668	-.029
A_Align_Bpractices	-.230	.190	-.125	-.107	-.009	.869
P_Time	.109	.008	.221	.226	.500	.655

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
I_Assignments	.819	.089	-.044	.123	-.192
I_Activities	.811	.232	.035	.228	-.256
I_Assess	.808	.069	.226	.186	.148
I_St_Relevance	.753	.174	.256	-.244	.331
I_Performance	.682	.197	.273	.170	-.252
I_Align_St	.618	.507	.180	.217	.011
P_Rubrics	.358	.327	.290	.344	.126
A_edTPA2Mission	.218	.857	.132	-.057	.186
A_Mission2edTPA	.041	.838	.038	.230	.087
P_Performance	.411	.708	.085	.144	-.072
A_ImprovedRep	.186	.005	.909	.158	.064
A_Reputation	-.012	.374	.831	.119	-.121
A_MissionChange	.547	-.034	.702	.074	-.075
P_Tasks	.131	.078	.007	.804	.008
P_Curriculum	-.002	.554	.093	.651	-.105
P_Time	.132	.024	.217	.606	.584
P_PDActivities	.259	.134	.334	.582	.077
A_Align_Bpractices	-.233	.150	-.139	.035	.840

Rotated Component Matrix[a]

	Component			
	1	2	3	4
I_Activities	.856	.223	.199	.017
I_Assignments	.840	.100	.112	-.052
I_Performance	.733	.179	.147	.254
I_Assess	.730	.124	.272	.240
I_St_Relevance	.610	.268	-.099	.313
I_Align_St	.590	.523	.251	.181
A_Align_Bpractices	-.479	.269	.229	-.078
A_edTPA2Mission	.147	.884	.005	.161
A_Mission2edTPA	.016	.832	.243	.037
P_Performance	.416	.701	.143	.083
P_Tasks	.144	.055	.784	-.044
P_Time	-.035	.095	.750	.223
P_PDActivities	.241	.129	.609	.304
P_Curriculum	.047	.502	.602	.045
P_Rubrics	.312	.341	.395	.285
A_ImprovedRep	.169	-.002	.219	.902
A_Reputation	.034	.325	.118	.811
A_MissionChange	.551	-.034	.113	.697

Rotated Component Matrix[a]

	Component		
	1	2	3
I_Activities	.834	.268	.090
I_Assignments	.810	.136	-.005
I_Performance	.746	.162	.264
I_Assess	.719	.168	.325
I_St_Relevance	.678	.127	.170
I_Align_St	.621	.526	.235
A_Align_Bpractices	-.481	.355	.051
A_Mission2edTPA	.075	.831	.098
A_edTPA2Mission	.257	.751	.062
P_Curriculum	.022	.689	.330
P_Performance	.471	.660	.083
A_ImprovedRep	.248	-.089	.864
A_Reputation	.153	.177	.709
A_MissionChange	.606	-.126	.631
P_Time	-.099	.356	.595
P_PDActivities	.204	.308	.575
P_Rubrics	.323	.409	.421
P_Tasks	.033	.386	.393

The 3-factor model only accounted for 60% of the variance of the loadings within factors, but the items converged nicely on the constructs, increasing the likeliness of evidence for convergence validity (Allen & Yen, 1979). The four and five-factor models accounted for 67% and 74% of the variance, respectively.

Exploratory Factor Analysis Results for the Three-Factor Model.

The correlations between items ranged from -.01 to .81 (see Table 4.8, "Correlation Matrix"). The Kaiser-Meyer-Olkin measure of sampling adequacy yielded a value of .71, and Bartlett's test of sphericity (df, 153) was significant, $p<.01$, indicating that the assumption of sphericity was not met, and that the Geisser-Greenhouse test should be interpreted. The factor analysis was held constant to 3 constructs. The three constructs explained 60% of the variance (see Table A. 2. Factor Analysis, below). The items loading into each construct were grouped in order to label each of the three constructs (see Table A. 3 "Rotated Component Matrix", below).

Table 6.2. Factor Analysis - 3 Factors

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.765	37.582	37.582	6.765	37.582	37.582	4.404	24.468	24.468
2	2.295	12.748	50.329	2.295	12.748	50.329	3.322	18.458	42.926
3	1.732	9.622	59.951	1.732	9.622	59.951	3.064	17.025	59.951
4	1.343	7.461	67.412						
5	1.265	7.029	74.441						
6	1.001	5.561	80.001						
7	.627	3.481	83.482						
8	.553	3.071	86.553						
9	.519	2.883	89.437						
10	.366	2.034	91.471						
11	.341	1.894	93.366						
12	.294	1.634	94.999						
13	.261	1.451	96.451						
14	.227	1.263	97.714						
15	.164	.911	98.626						
16	.115	.639	99.265						
17	.080	.444	99.708						
18	.053	.292	100.00						

Table 6.3. Rotated Component Matrix

	Component		
	1	2	3
I_Activities	.834	.268	.090
I_Assignments	.810	.136	-.005
I_Performance	.746	.162	.264
I_Assess	.719	.168	.325
I_St_Relevance	.678	.127	.170
I_Align_St	.621	.526	.235
A_Align_Bpractices	-.481	.355	.051
A_Mission2edTPA	.075	.831	.098
A_edTPA2Mission	.257	.751	.062
P_Curriculum	.022	.689	.330
P_Performance	.471	.660	.083
A_ImprovedRep	.248	-.089	.864
A_Reputation	.153	.177	.709
A_MissionChange	.606	-.126	.631
P_Time	-.099	.356	.595
P_PDActivities	.204	.308	.575
P_Rubrics	.323	.409	.421
P_Tasks	.033	.386	.393

The first factor overlapped perfectly with the a priori 6-item Instruction scale. Therefore, the construct was interpreted as Instructional Alignment, because the questionnaire items that loaded onto this construct involved using standards during instructional activities to align to the edTPA.

The second factor contained the same items as the Mission Alignment construct from the 6-factor model, but also included the "Curriculum" item (see item 4 below). Factor 2 was interpreted as Mission Alignment, because its items addressed perceptions of whether the edTPA and the preparation programs' mission are aligned, and supported with self-evaluation of performance: 1) My program's mission is consistent with the expectations of the edTPA; 2) The demands of the edTPA are consistent with and support my program's mission; 3) My teacher preparation program uses the edTPA scores as one data point to internally evaluate our own performance; 4) The program for which I teach has arranged its curriculum to prepare candidates for the edTPA implementation. The Mission Alignment construct (factor) explained 18% of the variance.

The third factor contained the same items as factors 3 and 4 from the 6-factor model, and also included items from factors 5 and 6 of the 6-factor model. This third factor could be interpreted as Program Improvement, as its items refer to perceived changes that the teacher preparation programs made to align to the edTPA: 1) The edTPA has improved the reputation of my program; 2) My program's reputation has strengthened since the edTPA's implementation; 3) My program's mission has changed as a result of the implementation of the edTPA; 4) Staff and faculty time is allocated so we can assist with one or more edTPA-related procedures; 5) My program uses the edTPA scores to select professional development activities to guide faculty practice; and 6) My teacher preparation program incorporates the edTPA rubric concepts in activities to guide faculty practice. This program improvement model explained 17% of the variance, and supports Peck et al.'s (2009) argument that program-wide,

structured opportunities that allow faculty to interpret data can lead to individual learning and program improvement.

Exploratory Factor Analysis Results for the Four-Factor Model.

The correlations between items ranged from -.01 to .81 (see Table 4.8, "Correlation Matrix"). The Kaiser-Meyer-Olkin measure of sampling adequacy yielded a value of .71, and Bartlett's test of sphericity (df, 153) was significant, $p<.01$, indicating that the assumption of sphericity was not met, and that the Geisser-Greenhouse test should be interpreted. The factor analysis was held constant to 4 constructs. The four constructs explained 67% of the variance (see Table A. 4. Factor Analysis - $-Factor Model, below). The items loading into each construct were grouped in order to label each of the four constructs (see Table A. 5 "Rotated Component Matrix", below).

Table 6.4 Factor Analysis – 4-Factor Model

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.765	37.582	37.582	6.765	37.582	37.582	4.167	23.148	23.148
2	2.295	12.748	50.329	2.295	12.748	50.329	2.994	16.633	39.782
3	1.732	9.622	59.951	1.732	9.622	59.951	2.493	13.852	53.634
4	1.343	7.461	67.412	1.343	7.461	67.412	2.480	13.778	67.412
5	1.265	7.029	74.441						
6	1.001	5.561	80.001						
7	.627	3.481	83.482						
8	.553	3.071	86.553						
9	.519	2.883	89.437						
10	.366	2.034	91.471						
11	.341	1.894	93.366						
12	.294	1.634	94.999						
13	.261	1.451	96.451						
14	.227	1.263	97.714						
15	.164	.911	98.626						
16	.115	.639	99.265						
17	.080	.444	99.708						
18	.053	.292	100.00						

Table 6.5 Rotated Component Matrix – 4-Factor Model

	Component			
	1	2	3	4
I_Activities	.856	.223	.199	.017
I_Assignments	.840	.100	.112	-.052
I_Performance	.733	.179	.147	.254
I_Assess	.730	.124	.272	.240
I_St_Relevance	.610	.268	-.099	.313
I_Align_St	.590	.523	.251	.181
A_Align_Bpractices	-.479	.269	.229	-.078
A_edTPA2Mission	.147	.884	.005	.161
A_Mission2edTPA	.016	.832	.243	.037
P_Performance	.416	.701	.143	.083
P_Tasks	.144	.055	.784	-.044
P_Time	-.035	.095	.750	.223
P_PDActivities	.241	.129	.609	.304
P_Curriculum	.047	.502	.602	.045
P_Rubrics	.312	.341	.395	.285
A_ImprovedRep	.169	-.002	.219	.902
A_Reputation	.034	.325	.118	.811
A_MissionChange	.551	-.034	.113	.697

Again, the first factor overlapped perfectly with the a priori and 6-factor Instruction scale's items. Therefore, the construct was interpreted as Instructional Alignment, because the questionnaire items that loaded onto this construct involved using standards during instructional activities to align to the edTPA.

The second factor also overlapped perfectly with the 6-factor Mission Alignment scale's items. Therefore, the second factor was interpreted as "Mission Alignment" because its items addressed perceptions of whether the edTPA and the preparation programs' mission are aligned, and supported with self-evaluation of performance: 1) My program's mission is consistent with the expectations of the edTPA; 2) The demands of the edTPA are consistent with and support my program's mission; and 3) My teacher preparation program uses the edTPA scores as one data point to internally evaluate our own performance. The Mission Alignment construct (factor) explained 17% of the variance.

Factor 3 included items from constructs four, five, and six of the six-factor model. This construct was interpreted as Program Improvement, because the items describe opportunities that the programs may have provided in order to guide faculty practice: 1) My program has changed its processes to incorporate the tasks in the edTP; 2) Staff and faculty time is allocated so we can assist with one or more edTPA-related procedure; 3) My program uses the edTPA scores to select professional development activities to guide faculty practice; 4) The program for which I teach has arranged its curriculum to prepare candidates for the edTPA implementation. This third factor explained 14% of the variance.

Factor 4 overlapped perfectly with the "Program Reputation Change" construct in the 6-factor construct. Therefore, factor 4 was interpreted as Program Reputation Change, because its items addressed perceptions of whether the preparation program's reputation had changed since it began responding to the edTPA adoption: 1) The edTPA has improved the reputation of my program; 2) My program's reputation has

strengthened since the edTPA's implementation; 3) My program's mission has changed as a result of the implementation of the edTPA. The Program Reputation Change construct explained 14% of the variance.

6.2 Email Questionnaire Request and Consent

Please fill out the 15-minute dissertation survey (link below) for your chance to win a $40 Starbucks gift card.

As one of the initial teacher education programs to pilot the edTPA, you may be interested in the status of program and instructional practice across Washington State. Your teacher education program was chosen to participate in this survey because of its history in leading program improvement strategies in Washington State. If your work involves teacher education, participating in this 22-question study will further inform your practice, as well as the practice of teacher educators across Washington State.

And you could win one of four $40 gift cards just for filling out a 15 minute survey!

Please read the consent paragraph below. Clicking on the link to the survey indicates your consent to participate in this dissertation research study.

Thank you for contributing to information regarding the edTPA, and for contributing to this dissertation research.

https://www.surveymonkey.com/r/edTPA_Survey

Maria Muzzo

Ph.D. Candidate

University of Washington

---

https://www.surveymonkey.com/r/edTPA_Survey

The purpose of this study is to determine how, if at all, the edTPA implementation has influenced the nature of practice of teacher preparation programs in Washington State. This survey will focus on the edTPA implementation within your program. The survey contains 22 statements and should take no more than 15 minutes. You are free not to answer any questions you do not wish to answer. The survey will contain no links between email address and survey responses. Email addresses will be collected for the sole purpose of choosing gift card winners. Although program directors could ultimately learn the results of this study through reading the dissertation, publication, or report, they will not know who has or has not participated. Names will not be used in reporting the results of the study. Results will be combined across institutions according to roles, so that individuals within any given institution are not identifiable.

We hope the results of this study will help teacher educators understand the implementation of evidence-based policies like edTPA. We intend this study to inform policy makers and practitioners concerned with the quality of teacher education. No direct, personal benefit can be expected from study procedures.

CLICK ON THE LINK BELOW TO TAKE THE SURVEY

https://www.surveymonkey.com/r/edTPA_Survey

Best regards,

Maria Muzzo

Ph.D. Candidate

University of Washington