

Lindenwood University

Digital Commons@Lindenwood University

Dissertations

Theses & Dissertations

Summer 7-2015

A Mixed-Method Investigation of the Missouri Pre-Service Teacher Assessment Pilot Program at a Private Midwestern University

Robyne Elder
relder@lindenwood.edu

Follow this and additional works at: <https://digitalcommons.lindenwood.edu/dissertations>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Elder, Robyne, "A Mixed-Method Investigation of the Missouri Pre-Service Teacher Assessment Pilot Program at a Private Midwestern University" (2015). *Dissertations*. 338.
<https://digitalcommons.lindenwood.edu/dissertations/338>

This Dissertation is brought to you for free and open access by the Theses & Dissertations at Digital Commons@Lindenwood University. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital Commons@Lindenwood University. For more information, please contact phuffman@lindenwood.edu.

A Mixed-Method Investigation of the Missouri Pre-Service Teacher Assessment Pilot
Program at a Private Midwestern University

by

Robyne Elder

A Dissertation submitted to the Education Faculty of Lindenwood University

in partial fulfillment of the requirements for the

degree of


Doctor of Education

School of Education

A Mixed-Method Investigation of the Missouri Pre-Service Teacher Assessment Pilot
Program at a Private Midwestern University

by
Robyne Elder


This dissertation has been approved in partial fulfillment of the requirements for the
degree of
Doctor of Education
at Lindenwood University by the School of Education



Dr. Lynda Leavitt, Dissertation Chair

7-14-15

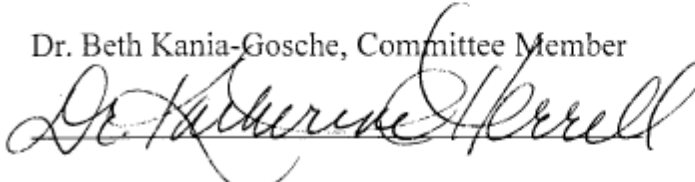
Date



Dr. Beth Kania-Gosche, Committee Member

7-14-15

Date



Dr. Katherine Herrell, Committee Member

7/14/2015

Date

Declaration of Originality

I do hereby declare and attest to the fact that this is an original study based solely upon my own scholarly work here at Lindenwood University and that I have not submitted it for any other college or university course or degree here or elsewhere.

Full Legal Name: Robyne Lynn Elder

Signature: Robyne Lynn Elder Date: 7/14/15

Acknowledgements

I am truly thankful for the support and inspiration from professors, teachers, friends, and family. To Dr. Beth Kania-Gosche for giving me the opportunity to pursue my doctorate and acting as the driving force for this topic and accompanying data, and to Dr. Lynda Leavitt for her constant feedback, guidance, and time. They are both women I aspire to be and I thank them for their service to research and education. I am also thankful for the conversations and classes with Dr. Terry Stewart and Dr. Graham Weir, experts in the world of leadership; I am grateful for all they have taught me regarding what it means to truly lead. My committee member and a student teacher supervisor, Dr. Kate Herrell, has given her time to work with me in different capacities and I appreciate this as well as her and other faculty members and adjunct instructors who are so vital to the student teaching experience. Moreover to Dr. Kevin Winslow, my statistics professor and office mate, whose teaching style inspired me to understand and appreciate quantitative statistics; and I thank him for his kindness and company during my days at Lindenwood.

Of course, none of this would be possible without my family. I want to thank my parents for showing me the importance of work ethic and education, and supporting me throughout all my endeavors in life. My sister has been a role model since I was young, and I thank her for introducing me to the world of education. Finally to my amazing family, Brian, Will, and Charlie, they are my heart and I thank them for their constant support and inspiration.

Abstract

In order to evaluate the teacher education program for the state of Missouri, the researcher investigated the piloted MoPTA at a private Midwestern university. Through evaluating the piloted MoPTA program, this study aimed to address possible changes needed within the teacher education program itself to better prepare future educators not only for the assessment, but more importantly, the classroom. In order to evaluate the program, the researcher observed scoring sessions for Tasks 1-4 and analyzed feedback from the evaluators of the tasks (university supervisors and faculty) in the fall of 2014 and the spring of 2015. Furthermore, this study examined the scores received from each task during the piloted school year (fall of 2014 and spring of 2015). The researcher analyzed the scores for the following comparisons: Tasks 1-4 (i.e. Task 1 overall scores to Task 2 overall scores); undergraduate students with graduate students' scores; inter-rater reliability (comparing the scores of multiple raters for one student); and K-12, elementary, middle, and secondary teacher candidates. By completing quantitative analyses of the comparisons through examining approximately 276 teacher candidates' scores, the researcher ascertained: student performance on Tasks 2 and 3 was significantly lower than Task 4; undergraduate students performed lower on Tasks 1 and 3 in comparison to graduate students; inter-rater reliability had a low correlation for Tasks 1, 3, and 4, but Task 2 reported a high correlation; and there were no differences between elementary teacher candidates and secondary/K-12 teacher candidates. Through analysis of qualitative data the researcher ascertained that the university supervisors and faculty found the scoring sessions for MoPTA helpful and that changes to the university's curriculum were necessary to better prepare teacher candidates. The researcher suggests

adding more data analysis, critical thinking, and writing courses at the university would better prepare teacher candidates; and ongoing scoring sessions and further professional development regarding changes in MoPTA and inter-rater reliability would benefit teacher candidates and consistency among university supervisors and faculty.

Table of Contents

List of Tables	viii
List of Figures	ix
Chapter One: Introduction	1
Background of the Study	1
Purpose of the Study	4
Rationale	8
Research Questions	11
Null Hypotheses	11
Limitations	12
Definition of Terms	14
Summary	17
Chapter Two: The Literature Review	18
Introduction	18
Organization of the Literature Review	19
Missouri Pre-Service Teacher Assessment (MoPTA)	19
Task 1	22
Task 2	24
Task 3	26
Task 4	28
edTPA	30
California’s Teacher Preparation Assessments	34
Aligning TEPs to Standards	36

Online Feedback and Foliotek	39
Benefits to TPAs	41
Collaboration.....	41
Critical thinking skills.....	43
Providing and gathering data.	44
Disadvantages to TPAs	45
Cost.	45
Workload.....	48
Teacher attrition.	53
Suggestions for Teacher Recruitment and Retention.....	54
Summary	57
Chapter Three: Methodology	59
Purpose.....	59
Surveys.....	60
Scores	61
Methodology	61
Hypotheses	63
Research Questions.....	64
Limitations	65
Teacher Preparation Assessment’s Reliability and Validity.....	67
Inter-rater reliability.....	67
Purpose.....	69
The Research Site and Participants.....	70

Summary	71
Chapter Four: Results	73
Overview	73
Null Hypothesis 1	73
Null Hypothesis 2	77
Null Hypothesis 3	79
Null Hypothesis 4	83
Research Question 1	85
Research Question 2	91
Research Question 3 and Research Question 4.....	93
Research Question 5	98
Summary	99
Chapter Five: Discussion and Reflection.....	101
Overview.....	101
Discussion.....	102
Null Hypothesis 1.	102
Null Hypothesis 2.	104
Null Hypothesis 3.	106
Null Hypothesis 4.	107
Participants’ perceptions of scoring sessions and evaluating MoPTA online via Foliotek.	108
Changing content of lessons after scoring MoPTA.	109
Participants’ perception of teacher candidate preparation.	110

Participants' perception of full implementation of MoPTA.....	110
Reflection on the Piloted Program.....	111
Recommendations for Program.....	114
Recommendations for Future Research.....	118
Conclusion.....	119
Appendix A: Survey Questions.....	131
Appendix B: Permission to use Survey and the Study Site.....	132
Appendix C: Observational Data from MoPTA Scoring Sessions.....	133
Appendix D: NIH Certificate.....	134
Appendix E: Informed Consent.....	135
Vitae.....	136

List of Tables

Table 1. MoPTA Pilot Timeline for Study University.....	4
Table 2. MoPTA Summary.....	21
Table 3. CalTPA Inter-Rater Reliability Coefficient.....	68
Table 4. Results of Scores from Tasks 1-4: Fall 2014.....	74
Table 5. Results of Scores from Tasks 1-4: Spring 2014	74
Table 6. Results from ANOVA Test for Tasks 1-4: Fall 2014.....	75
Table 7. Results from ANOVA Test for Tasks 1-4: Spring 2014	75
Table 8. Scheffe Test: Tasks 1-4 Fall 2014	76
Table 9. Scheffe Test: Tasks 1-4 Spring 2015	76
Table 10. BAT v. MAT Students.....	78
Table 11. Scheffe Test: BA v. MA Students	78
Table 12. Tasks 1-4: Elementary v. Secondary/K-12 Students	85
Table 13. Scheffe Test: Elementary v. Secondary/K-12 Students.....	85

List of Figures

Figure 1. Foliotek screen shot.....	41
Figure 2. Task 1 rubric.....	62
Figure 3. Inter-rater reliability for Task 1.....	80
Figure 4. Inter-rater reliability for Task 2.....	81
Figure 5. Inter-rater reliability for Task 3.....	82
Figure 6. Inter-rater reliability for Task 4.....	83

Chapter One: Introduction

Background of the Study

In 2003, the United States Department of Education issued a report urging schools to hire teachers of the highest quality. In their definition of “highly qualified” they referred to the No Child Left Behind Act of 2001 which clarified components teachers must possess: “a bachelor’s degree...state certification...and competence in their subject area” (U.S. Department of Education [USDOE], 2003, p. 4). USDOE (2003) clarified elementary teachers show this competency through state tests, and middle and high school teachers exhibit skill through “passing a rigorous exam of their content knowledge; majoring in their subject as an undergraduate; earning a graduate degree in their subject; accumulating the coursework equivalent to an undergraduate major; or attaining an advanced certificate or credential” (p. 4). However, at the time of this writing, 12 years after this USDOE report, these measures showed no correlation with student achievement (Caughlan & Jiang, 2014). States have used various tests, as mentioned by USDOE as “state tests” or “rigorous exam(s)” to measure teacher performance in order to gain certification and show competency. Forty states have used the testing company Educational Testing Service (2014) to supply these types of assessments. For example, states such as Kansas and New Hampshire use the ETS supplied *Praxis II* exam as requirement to receive certification; however, Missouri moved from the *Praxis II* to a different ETS created assessment. ETS partnered with the Missouri Department of Elementary and Secondary Education (MODESE) to create the assessments under the title of the Missouri Educator Gateway Assessments (MEGA). The exam for certification was known as the Missouri Content Assessments and it started in

the fall of 2014 (ETS, 2014; MEGA, 2015b). Although this type of standardized assessments supplied by states is popular, a new type of assessment began to emerge to validate competency beyond a paper and pencil or computer-based exam. Caughlan and Jiang (2014) stated it was more valid to investigate the measures of teacher quality by means of an assessment that examined teacher performance during their student teaching experience. This type of pre-service teacher assessment earned national attention as it progressed from New York to California to Missouri. The two organizations, National Council for Accreditation of Teacher Education (NCATE) and Teacher Education Accreditation Council (TEAC) required universities show documentation of the teacher candidate's knowledge of a range of topics in education and the implementation of this knowledge in the classroom during the student teaching experience (Caughlan & Jiang, 2014). This type of assessment was referenced by various names in different states, but the overall term for the evolving way to ascertain if teachers were ready for the classroom was TPAs or Teacher Performance Assessments (Caughlan & Jiang, 2014). The national TPA was titled edTPA and at the time of this writing was fully implemented in 13 states, with 20 states exploring the assessment (American Association of Colleges for Teacher Education [AACTE], 2015), para. 1). The other 17 states chose different paths, such as creating their own assessment, i.e. California (CalTPA) and Missouri (MoPTA). This study aimed to examine one TPA in Missouri, the Missouri Pre-Service Teacher Assessment, during its pilot school year in a private university setting.

The transition to this Teacher Performance Assessment, called MoPTA for the Missouri Pre-Service Teacher Assessment, aligned with the new certification requirements, which aligned to the Missouri Standards for the Preparation of Educators

(MoSPE) standards, as outlined by the Missouri Department of Elementary and Secondary Education (n.d.a.). They stated the following steps toward becoming an educator: “Grade Point Average, Missouri Educator Profile, Missouri General Education Assessments, Missouri Content Area Exams, Missouri Performance Assessments, and Certification Requirements” (MODESE, n.d.a., para. 1). As stated in a memo from the Director of Education Preparation, starting in the fall of 2013, students needed a GPA of 2.75 (cumulative), 3.0 for content and professional education in order to be eligible for certification (Hariston, 2013, para. 2). The Missouri Educator Profile, or the MEP, is a web-based evaluation of “work-style preferences used to support the development of effective educator work habits” which must be taken during the student’s enrollment in the Teacher Education Program to gain certification (MEGA, 2015a, para. 1). This is not a grade-based assessment, but a method future educators can use to understand their own occupational behaviors (MEGAa, 2015). Furthermore, students must take and pass the Missouri General Education Assessment (MoGEA), which started in September, 2013 when MODESE replaced the previous College Base (CBASE) exam with the MoGEA (MEGA, 2015b). As stated earlier, Missouri replaced the *Praxis* exam with the MEGA assessment known as the Missouri Content Assessments (MEGA, 2015b). Students who had taken and passed the exam prior to August 31, 2014 did not need to take the Missouri Content Assessments; the implementation of the new exam began on September 2, 2014 (MODESE, n.d.a.). Another focus of change for Missouri requirements is the focus of this study, the completion of the Missouri Performance Assessments or MoPTA; which was implemented in September of 2014 (MODESE, n.d.a.). Once all the above are achieved, future educators completed certification requirements and should apply for said

certificate (MODESE, n.d.a.). The reason for the above changes university educator preparation programs in the state of Missouri was to create a consistency among the programs so the state can verify they are effectively training educators for the future; and to evaluate the university's programs are adhering to state guidelines (MODESE, 2014). In 2014, MODESE, via a webinar, instructed universities of the timeline they should work from when implementing the new certification requirements which clarified that by fall semester of 2017 all educator preparation programs were adhering to the new certification requirements (MODESE, 2014). This study aims to examine the need for such changes in educator preparation programs and the implementation of one facet of such programs, the performance assessment or for Missouri, the MoPTA. In order to accomplish this study, the researcher examined the pilot program of the MoPTA.

Purpose of the Study

The purpose of this study was to investigate the Missouri Pre-Service Teacher Assessment (MoPTA) piloted program evaluating Missouri teacher candidates at a private Midwestern university. The university established a pilot of MoPTA based on the new certification requirements from MODESE (n.d.a.) and adhered to the below timeline.

Table 1.

MoPTA Pilot Timeline for Study University

Event	Date
Volunteer pilot	Fall 2013
Volunteer scoring	Spring 2014
Original implementation	Fall 2014 (delayed until fall 2015)*
University pilot	2014-2015
Alternate Task 4 pilot	Spring 2015**
Alternate Task 4 scoring	May 2015**

Note. Due to policy issues regarding technology, full implementation was delayed one year (Hariston, 2014). To rectify the problem, MEGA (MODESE, 2015) established an alternate Task 4 artifact.

Through evaluating the piloted MoPTA program, this study aimed to address possible changes needed within the teacher education program itself in order to better prepare future educators not only for the assessment, but more importantly, the classroom.

Furthermore, the study hoped to provide possible steps to move forward with the use of the current TPA in terms of changes to be made or alternative assessments to be put in place. The MoPTA consists of four tasks for student teachers to complete during their student teaching experience (MEGA, 2014). The tasks covered the following topics: “Knowledge of Students and the Learning Environment” (MODESE, 2013b, para. 1); “Assessment and Data Collection to Measure and Inform Student Learning” (MODESE, 2013c, para. 1); “Designing Instruction for Student Learning” (MODESE, 2013d, para. 1); and “Implementing and Analyzing Instruction to Promote Student Learning” (MODESE, 2013e, para. 1). The tasks consisted of commentary on each question, located in textboxes, along with uploading required artifacts used as evidence to support the commentary (MEGA, 2014). The required artifacts ranged from contextual charts for Task 1 to a video component for Task 4. However, as noted in the timeline in Table 1, the video component for Task 4 was deemed optional and offered an alternative (Hariston, 2014; MODESE, 2015). The MoPTA tasks and artifacts are discussed more specifically in the review of literature located in Chapter Two. Once teacher candidates completed a task, university supervisors (full-time faculty, part-time faculty, and adjuncts serving as supervisors of teacher candidates during their student teaching practicum) and faculty (for the purposes of this dissertation, faculty will refer to faculty members at the study university who scored teacher candidates’ tasks but did not observe teacher candidates during their student teaching practicum) scored the tasks on a scale from 1-4, with 1 as

the lowest score and 4 as the highest. In order to evaluate the program, the researcher observed scoring sessions for Tasks 1-4 and analyzed feedback from the evaluators of the tasks (university supervisors and faculty) in the fall of 2014 and the spring of 2015. Furthermore, this study examined the scores received from each task during the piloted school year (fall of 2014 and spring of 2015). The researcher analyzed the scores for the following comparisons: Tasks 1-4 (i.e. Task 1 overall scores to Task 2 overall scores); undergraduate students with graduate students' scores; inter-rater reliability (comparing the scores of multiple raters for one student); and K-12, elementary, middle, and secondary teacher candidates. By completing quantitative analyses of the comparisons, the researcher hoped to accomplish the following: examine the task(s) where students excel, and where students struggle; analyze specific objectives not addressed in the university classroom; investigate discrepancies and lack thereof between scores of multiple raters per one student; and provide feedback regarding the performances of the undergraduate students versus the graduate students. Through this investigation of the piloted MoPTA, the researcher hoped to possibly pinpoint specific and necessary curriculum modifications to be implemented at the researched university to adhere to state and national standards for educators. Furthermore, the researcher found the results useful to other universities implementing like programs to compare this study's school to their own to anticipate possible areas of concerns and strengths. This study also holds merit as over 260 student teachers participated in the MoPTA.

Research has shown that the student teaching experience is vital to a pre-service teacher becoming a certified and highly effective instructor (Asplin & Marks, 2013). In the state of Missouri, this evolved into MoPTA or the Missouri Pre-Service Teacher

Assessment (MEGA, 2014), which was implemented in the fall of 2015. This assessment was divided into four tasks, completed by the teacher candidate to gain state certification; aligned with the state teacher standards as mandated by the Missouri Department of Elementary and Secondary Education (MODESE, 2013a). According to MODESE (2013b) the ability for each teacher candidate to meet the state standards, and therefore the tasks, informed the state of Missouri if the teacher candidate was ready to become a certified teacher.

In order to prepare and evaluate the teacher candidate for MoPTA, a private Midwestern university implemented a piloted MoPTA program in the summer of 2014 and continued into the following summer of 2015. This piloted program consisted of soon-to-be teacher candidates who completed the tasks and university supervisors and faculty who evaluated the tasks. During its full implementation the Educational Testing Service (ETS), the testing company, which partnered with MODSE to create MoPTA, would score the tasks (MEGA, 2014). However, university supervisors and faculty scoring the tasks served the purpose of not only preparing the teacher candidates for their teaching experience to begin the next semester, but it also served to train university supervisors regarding the tasks and evaluation of the teacher candidates using the new standards as set by MODESE (MEGA, 2014). Fernandez and Erbilgin (2009) explained the vital role university supervisors' play "in supporting student teachers' implementation of recent reforms and theories learned in coursework" (p. 94). By having the supervisors and faculty evaluate the completed tasks using the online portfolio system Foliotek (2014) they provided teacher candidates an overall score for each task and written feedback for the individual work of each textbox within the task. Furthermore, in the

researcher's experience, Foliotek provided students an opportunity to submit their work online as a way to organize their task submissions and receive scores and feedback in a timely manner. The Bill and Melinda Gates Foundation (2010) completed numerous research studies regarding the Measures of Effective Teaching (MET) and overwhelmingly found that using online submissions and gaining feedback promptly led to productive and exceptional teachers.

Each task required the teacher candidates to effectively communicate via the written word and supply evidence that the state standard was met and the task completed. According to a study completed by Lenhart, Arafah, Smith, and MacGill (2008) for the PEW Research Internet Project, 83% of teenagers' parents believed there was much more importance on writing proficiently than two decades ago (para. 8). The children of these parents were in complete agreement as 86% of them thought writing well was crucial to being successful in the present and future (para. 8). Furthermore, the same group of teens surveyed overwhelmingly agreed that using technology in writing motivated them to write well. According to this data and the congruent data from the MET Project (Bill and Melinda Gates Foundation, 2010) the researcher believed it was imperative for university supervisors to assist teacher candidates with their writing in terms of the MoPTA task completion. This study aimed to highlight skills of importance, such as writing, to be addressed in the university curriculum to show competency on performance assessments.

Rationale

MoPTA is a relatively untested teacher preparation program; therefore, this study was original in nature. However, many previous studies of similar programs around the country had been completed, most notably a study by Sandholtz and Shea (2012) that

examined the Performance Assessment for California Teachers (PACT). Many states, at the time of this study, implemented a state performance assessment as a national movement to create a more valid and reliable instrument for assessing teacher candidate performance opposed to student teachers simply evaluated by university supervisors' observations and evaluations alone (Sandholtz & Shea, 2012). As teaching evolved into a profession where instructors must go beyond simply telling students facts and ideas, the way teachers are assessed must change as well (Munby, Russell, & Martin, 2001). In order to accommodate this changing role of the educator, teacher assessments were created at the state level to address advanced and professional teaching practices (Tellez, 1996). Each state's test could be traced back to the Teacher Performance Assessment, which gave teacher preparation programs access to a multiple-measure assessment system aligned to state and national standards-including Common Core State Standards and the Interstate Teacher Assessment and Support Consortium (InTASC). This alignment served as a guide in the development of curriculum and practice around the common goal of making sure new teachers were able to teach each student effectively and improve student achievement (edTPA, n.d., para. 1).

At the time of this writing 35 states (including Washington, D.C.) implemented or explored the national TPA, edTPA with the other 16 states choosing a different path, i.e. establishing their own TPA (AACTE, 2015). In California it was named CalTPA and PACT and in Missouri referred to as the MoPTA, each one with its own specific tasks, artifacts, rubrics, and only system for scoring and uploading. The tests also come with a cost as they are serviced by testing companies. ETS partnered with MODESE and charged \$275; however, at the time of this writing an official cost was not established via

ETS (Missouri Association of Colleges for Teacher Education, 2014). Whereas edTPA partnered with Pearson and charged \$300 (edTPA, 2014, para. 1). Many advantages for pre-service teachers and their universities lie within the performance-based test. A pre-service teacher can be credited as an effective teacher not only by the university, but also via state and national standards (Sandholtz & Shea, 2012). Furthermore, due to an evaluation through an assessment and not solely through subjective observations and evaluations, a more valid and credible tool was sought through exams like MoPTA (edTPA, n.d.; Missouri Educator Gateway Assessments, 2014). Also, the assessments provided a specific and more accurate view of the educator's future as a teacher and offered opportunities to further education to improve any lacking skills before entering the education workforce (Sandholtz & Shea, 2012). Although a consistency among universities addressing state standards for teacher candidates across the state was a positive move for the field of education, there were concerns related to validity and consistent results across the board as well as agreement in scores provided to individual students (Sandholtz & Shea, 2012). One way MoPTA tried to rectify consistent results was through ETS scoring the summative tasks when it was fully implemented. During the pilot of MoPTA at the study university, the supervisors of student teachers and faculty scored the tasks.

The study conducted by Sandholtz and Shea (2012) investigated a comparison between “university supervisors’ predictions and teacher candidates’ scores on a teaching performance assessment” (p. 1). In their research they discussed a great discrepancy between university supervisors’ expectations and the actual results of the assessment. Sandholtz and Shea further discussed the results and revealed most university supervisors

either under-predicted or over-predicted the scores for the students they observed in the classroom. Although the study conducted by this researcher did not address such discrepancies, the Sandholtz and Shea research addressed gaps where this study hoped to complete.

Research Questions

Research Question 1: How do university supervisors perceive the process of evaluating teacher candidates' MoPTA tasks through the online portfolio system Foliotek?

Research Question 2: How did faculty change the content of their lessons after evaluating teacher candidates' completed task(s)?

Research Question 3: How do faculty and university supervisors perceive the teacher candidate preparation process (at this particular university)?

Research Question 4: After participating in the piloted program with accompanying training, how do faculty and university supervisors perceive teacher candidates' level of preparation for full MoPTA implementation in the fall of 2015?

Research Question 5: After participating in the piloted program with accompanying training, how do faculty and university supervisors perceive their preparation for full MoPTA implementation in the fall of 2015?

Null Hypotheses

Null Hypotheses 1: There will be no difference of scores between tasks (i.e. Task 1 to Task 2).

Null Hypotheses 2: There will be no difference in MoPTA scores between undergraduate teacher candidates and graduate teacher candidates.

Null Hypotheses 3: There will be no difference in MoPTA scores given for the same teacher candidate by university supervisors and faculty (i.e.: student A is given the same score by two different raters [university supervisor and/or faculty]).

Null Hypotheses 4: There will be no difference in MoPTA scores between K-12, early childhood and elementary, middle, and secondary teacher candidates.

Limitations

Inter-rater reliability was based on one student's scores from multiple raters, however, there were a few students who were only scored by one rater and due to the randomness of assigning the same task to multiple raters once or twice the same rater was given the same student's task to score twice. Although this happened a limited amount of times, it may influence a true inter-rater reliability. Teacher candidates were required to submit Task 1, but were then given one of the other tasks to complete (Task 2, 3, or 4), thus the scorer of the tasks were all given one or multiple Task 1s to score, and then one other task to score. The number of students given to each university supervisor and/or faculty member to score during the fall session for Task 1, were 58 scorers for 132 students; some scorers were given as few as two to score, whereas one was given 16 to score. This did not happen again, as Task 2 yielded 46 scorers for 51 students, Task 3 had 46 scorers for 53 students, Task 4 noted 37 scorers for 30 students; and for Tasks 2-4 some scorers were only given one student to score with a maximum of five students to score. Therefore, Task 1 for the fall may have produced invalid inter-rater results; however, when applied to the scores overall, the difference was minimal. Also, due to the amount of tasks given to each scorer to evaluate, the spring semester produced few multiple raters for one student's task; therefore, the researcher only measured inter-rater

reliability in the fall and not the spring. Furthermore, in regards to Null Hypothesis 2, for the fall MoPTA, students were required to declare the degree they were seeking (BA or MA) in Foliotek; however, this profile field was not consistently completed in Foliotek for the spring MoPTA, therefore, data was only compiled for the fall. Similar in nature, for Null Hypothesis 4, the data was only able to be compiled by elementary and secondary/K-12 for the spring due to student's indication of their certification in the spring, but not required for the fall. Finally, some tasks had an even number of textboxes, so if the score for each textbox ranged from 2-3 then the scorer had to gauge what score to give the task overall, a 2 or a 3. This would have been based on best judgment as opposed to calculated mean.

All surveys (see Appendix A) distributed to the university supervisors and/or faculty who participated in the scoring sessions were not returned. Out of the 60 surveys distributed during each scoring session, approximately 20 were returned. However, those that were returned included detailed feedback and this along with the PI's observational notes were enough to justify attributable qualitative data to the study. Also, there was not a Task 4 scoring session in the spring, due to availability of scorers and an understanding by the scorers with evaluating the task either at home or in their offices, therefore no surveys were distributed during that session.

Furthermore, the responses to the survey questions may have been impacted by various technology malfunctions during the start of each scoring session. These are outlined more clearly in the results section of the dissertation, but it was difficult to log in to the computers and access Foliotek with the appropriate username and password. Due to these delays, some surveys were not completed and some reflected negatively on the

portfolio system Foliotek because of these delays as opposed to assessing the system as a tool for MOPTA.

Finally, ETS made slight changes to MoPTA during the piloted program at the study university. Therefore, one version of the assessment was given in the fall and the second given in the spring. However, each task assessed the same area as well as each textbox; it was a matter of changing the individual textbox prompts to better address each task's objective. For example, Task 2.1 in the spring required inclusion and reflection on baseline data, but the fall Task 2.1 did not.

Definition of Terms

Calibration: a comparison between measurements, and as applied to MoPTA the rating of a 1 (lowest) to 4 (highest) for each task completed by the student teacher (Foliotek, 2014; Trucano, Swiler, Igusa, Oberkampf, & Pilch, 2006).

Cooperating Teacher: for the purposes of this study, the instructor in the school district where the teacher candidate (see definition below) was assigned. The teacher candidate taught and/or co-taught the cooperating teacher's class(es) and evaluated by the cooperating teacher (Private University, n.d.a.).

Educational Testing Service (ETS):

develops, administers and scores more than 50 million assessment tests annually in more than 180 countries, at more than 9,000 locations worldwide. In addition to assessments, we conduct educational research, analysis and policy studies and develop a variety of customized services and products for teacher certification, English language learning and elementary, secondary and postsecondary education. (ETS, 2014, para. 1)

Faculty: for the purposes of this dissertation, faculty will refer to faculty members at the study university who scored teacher candidates' tasks but did not observe teacher candidates during their student teaching practicum

Foliotek: an online portfolio system for teachers and students to upload a variety of documents; student teachers upload each MoPTA task (see definition below) to Foliotek (Foliotek, 2014).

Graduate Students: for the purposes of this study, graduate students will refer to students seeking a Master of Arts.

Measures of Effective Teaching (MET): launched by the Bill and Melinda Gates Foundation (2010) to try new ways to evaluate teaching strategies in order to produce a resource of effective teaching methods.

Missouri Department of Elementary and Secondary Education (MODESE): This department provides a variety of services to school districts, colleges, and other educational institutions in the state of Missouri. For the purposes of this dissertation, it will be used as a source for the requirements for student teachers in order to obtain a Missouri teaching certificate (MODESE, n.d.b.).

Missouri Pre-Service Teacher Assessment (MoPTA):

Aligned with Missouri's Teacher Standards and Quality Indicators (PDF), this test assesses the instructional capability of teacher candidates prior to receiving a Missouri state teaching license. It is designed to: develop more effective teachers in the classroom; identify strengths and areas for improvement of practice; contribute to a development plan for professional growth. The assessment consists

of four tasks, one formative and three summative. (Missouri Educator Gateway Assessments, 2014, para. 1)

Missouri Teacher Standards: A total of nine standards, which express the prospects for teachers in the state of Missouri. They are connected to the idea that instructors are empathetic and thoughtful experts in the field of education, who constantly search for innovative ways to inspire learning in all students (Missouri Department of Elementary and Secondary Education, 2013).

Reliability: obtaining the same results from a variety of tests (Rantanen, 2013).

Students' Evaluation of Teaching (SET): current student teachers give feedback and rate their own performance in the classroom (Rantanen, 2013).

MoPTA Tasks:

Task 1: “Knowledge of students and the learning environment” (MODESE, 2013b, para. 1). The task requires the teacher candidate to “demonstrate the knowledge and skills” which apply to his or her “understanding” of the “classroom in regard to...students, the school, and the community” (MODESE, 2013b, para. 1). Furthermore, the teacher candidate must discuss the “implications of these factors on instruction and student learning” (MODESE, 2013b, para. 1).

Task 2: “Assessment and data collection to measure and inform student learning” (MODESE, 2013c, para. 1). Teacher candidates are evaluated on the evidence from tests given to students and how the data from the assessments are used to increase student learning (MODESE, 2013c).

Task 3: “Designing instruction for student learning” (MODESE, 2013d, para. 1). The rubric evaluates the student teacher candidate on evidence, which shows theories,

learning strategies, and other effective instructional tools used to increase student learning (MODESE, 2013d).

Task 4: “Implementing and analyzing instruction to promote student learning” (MODESE, 2013e, para. 1). The student teacher coordinators are required to show proof of the planning of the lesson and to show how the standards aligned with the goals set for learning (MODESE, 2013e).

Teacher Candidate (TC): for the purposes of this study, the teacher candidate referred to the college student seeking teacher certification (Private University, n.d.a.).

Undergraduate Students: for the purposes of this study, undergraduate students will refer to students seeking a Bachelor of Arts.

University Supervisor: for the purposes of this study, the university supervisor will refer to the university employee evaluating the teacher candidate (Private University, n.d.a.). The job responsibilities include observing the teacher candidate four times, completing observation forms, communicating with the cooperating teacher and administration of the teacher candidate’s placement, and assisting the teacher candidate through the student teaching practicum.

Summary

The purpose of this study was to investigate the piloted MoPTA program at one private university. Future educators must be qualified for the classroom upon graduation and the researcher believed TEPs were crucial in producing this type of workforce (Chung, 2008; Sandholtz & Shea, 2012). The researcher also believed the aspects of effective TEPs was worthy of study. These topics were addressed in the next chapter within a review of the current literature.

Chapter Two: The Literature Review

Introduction

Fostering education majors to becoming qualified educators continues to be crucial to the success of K-12 students. As the number of undergraduate and graduate pre-service teachers graduate, they offer a pool of applicants for the many soon-to-be retired teachers leaving their schools (Dynarski, 2014). Taking this into consideration, it was clear this could mean more business for universities' TEPs. In 2014, "more than 2,000 teacher preparation programs graduate more than 200,000 students a year, which generates billions of dollars in tuition and fees for higher education institutions" (Dynarski, 2014, para. 1). Therefore it was imperative for universities to implement highly effective TEPs in order to train future educators, especially due to the negativity that surrounded education. According to Sawchuck (2014) from *Education Week*, the constant changes to the field along with state budget cuts, was diminishing the view of teaching as a job with longevity. Across the country, students enrolled in TEPs fell 10% from 2004 to 2012 (Sawchuck, 2014, para. 5). The state of California saw teacher training diminish by 53%, with other enrollments down across the nation as seen in states such as New York, Texas, and North Carolina with a decrease of 20% during the past three years (Westervelt, 2015, para. 1). It was clear that in order for TEPs to be successful, they must show they can produce effective educators who can weather the constant changes in the field and stay focused on teaching students using the methods and strategies gained from their university experience.

Organization of the Literature Review

The literature review will begin with a description of three different types of TPAs. The first is the focus of the study, Missouri's TPA (MoPTA), then a comparison to the national program, edTPA, ending with another perspective of an individual state test via California's TPA (CalTPA). The review then continues with an account of the use of standards to align TPAs, such as National Standards for edTPA, and MoSPE standards for MoPTA. Following the standards is a description of how such TPAs are scored online and the feedback teacher candidates are likely to receive. Benefits to TPAs, such as collaboration, critical thinking skills, and access to data are described followed by disadvantages, i.e. cost, workload, and teacher attrition; however, following the analysis of attrition are research-based recommendations for TPAs to recruit and keep teachers in TEPs.

Missouri Pre-Service Teacher Assessment (MoPTA)

According to the Missouri Educator Gateway Assessments (MEGA, 2014) website, the Missouri Department of Elementary and Secondary Education (MODESE) and the Educational Testing Service (ETS) joined to establish an assessment for the state of Missouri which was standards-based. The purpose was to determine the teacher candidate's performance level in coursework dealing with the content of the subject matter as well as the student teaching practicum (MEGA, 2014). It was also established to: "develop more effective teachers in the classroom; identify strengths and areas for improvement of practice; and contribute to a development plan for professional growth" (MEGA, 2014, para. 2). This was in alignment with other established pre-service assessments such as edTPA (n.d.), which cost \$300 (edTPA, 2014, para. 1) in comparison

to MoPTA at a cost of \$275. This not only was used to evaluate teacher candidates, but other assessments were established to evaluate principals, counselors, and librarians. The MoPTA was a requirement for certification by the state of Missouri starting in the fall of 2017 (MODESE, 2014). The assessment had to be accomplished during the student teaching practicum and completely finished by the time student teaching was completed (MEGA, 2014). For teacher candidates, the MoPTA consisted of four tasks and each required written explanation and uploading evidence, which supported the written work. For example, Task 1 required a contextual chart and Task 2 required an example of baseline data; Table 2 explains the artifacts required for each task. Task 4 required the teacher candidate to upload a video as an artifact; however, due to technology policies in some host schools for teacher candidates, MEGA established a Task 4 alternate artifact (Hariston, 2014; MODESE, 2015). All documents required for the tasks had to be submitted online. In this study they were submitted via the online portfolio system, Foliotek. However, universities and/or ETS may have used a different system when it was implemented in the fall of 2015. Table 2 outlines the requirements for each task; specific requirements are listed in the sections after Table 2.

Table 2.

MOPTA Summary

Task	Artifacts	Summary of Written Response Prompts
1: Knowledge of Students and the Learning Environment	Contextual Factors Chart Instructional and Support Resources Chart Completed Student Interest Inventory Method of Introduction to parents	Select factor from each row of Contextual Factors chart that will most impact student learning, identify instructional strategy and learning activity with rationale Write about two resources, select a third and connect to CF chart Analyze student interest inventory to support student engagement Link method of communication to classroom demographics Analyze classroom management and technology rules/procedures
2: Assessment and Data Collection to Measure and Inform Student Learning (Reading for elementary)	A baseline data document Two representative pages of assessment Scoring guide or rubric Graphic representation of class assessment data Work samples from two focus students	Align assessment to standards, justify method and data collection What evidence of student learning will this assessment provide? Differentiate/modify the assessment for two focus students Analyze data, share with class and two focus students Reflect on the assessment, weaknesses, how will results inform your future teaching How did the two focus students do? Reflect on modifications
3: Designing Instruction for Student Learning (Math for elementary)	Lesson plan Modification/adaptions for two focus students Instructional artifact Representative student work sample (not one of the focus students) Two focus students' work samples	Link lesson plan to learning theory, standards, previous content. Three instructional strategies w/rationale, connect to learning goal Learning activities informed by student strengths, needs, and class demographics Materials and resources used (technology must be included) Two focus students with adaptations/modifications for each, reflection on the lesson for each focus student

Table 2. Continued.

		Analyze the lesson: feedback, meaningful learning, interaction, engagement Reflect on lesson, future teaching, what will you do for those who did not achieve learning goals?
4: Implementing and Analyzing Instruction to Promote Student Learning	15 minute video* (or three, 5 minute segments) Lesson Plan Two focus students' work samples	Align lesson to standards and student's learning needs Instructional strategies to engage students in academic language, critical thinking, inquiry, and integration of reading into content area (rationale) These should also be in the video* Cite evidence from video* of classroom management, verbal and nonverbal communication techniques Reflect on lesson for two focus students and the whole class

Note. MEGA, 2014. * However, due to technology policies in some host schools for teacher candidates, MEGA established a Task 4 alternate artifact (Hariston, 2014; MODESE, 2015).

Task 1.

The first task required by MoPTA regarded “Knowledge of Students and the Learning Environment” in that it requested the teacher candidate to “demonstrate the knowledge and skills that pertain to your understanding of the context of your classroom in regard to your students, the school, and the community; and you will identify implications of these factors on instruction and student learning” (MODESE, 2013b, para. 1). This task addressed various quality indicators within Missouri state standards 2, 3, 4, 5, 6, 8, and 9 (MODESE, 2013a). The task was broken down into various textboxes addressing different facets of working with students' prior knowledge and creating the most effective learning environment (MODESE, 2013b). The writing could not be more than 21,000 characters and had to follow these guidelines: responded to each prompt

within the task; referenced the uploaded evidence or artifacts to support the writing; defined, evaluated, and reflected on the evidence; and “reflects in what ways the evidence you have collected impacts your understanding of the knowledge of students and the classroom learning environment” (MODESE, 2013b, para. 2). The artifacts submitted were not by the teacher candidate’s discretion but instead were mandated by MODESE. For Task 1 the required evidence was “the Contextual Factors Chart; the Instructional and Support Resources Chart; one completed student interest inventory; and a form of introduction” (MODESE, 2013b, para. 2). The two charts were developed by MODESE and ETS and placed on their website as well as the Missouri Educator Gateway Assessments’ website for teacher candidates ease of access. Those exact charts had to be used in the construction of Task 1; however, teacher candidates were allowed to create their own student interest survey (questions inquiring students various interests in order to guide lesson planning) and a form of introduction (i.e. letter home to parents, introductory email sent prior to school starting), or student use of the templates posted to MODESE’s and MEGA’s websites (MODESE, 2013b). Below outlines the requirements for textbox 1 of Task 1 in order to provide an example of the type of prompts the students would respond to throughout the task:

Textbox 1.1.1: Community, District, School Contextual Factors that Influence

Instruction: A. Based on your chosen community factor, identify and describe one possible instructional strategy and one learning activity that you could use in your classroom to further student learning. Provide a rationale for choosing that strategy and activity and explain why it appropriately connects to your chosen factor. B. Based on your chosen district factor, identify and describe one possible

instructional strategy and one learning activity that you could use in your classroom to further student learning. Provide a rationale for choosing that strategy and activity and explain why it appropriately connects to your chosen factor. C. Based on your chosen school factor, identify and describe one possible instructional strategy and one learning activity that you could use in your classroom to further student learning. Provide a rationale for choosing that strategy and activity and explain why it appropriately connects to your chosen factor. (MODESE, 2013b, para. 3)

This task was designed to be formative in nature in that it was to be completed during the first few weeks of student teaching and to be used as a tool to plan future lesson plans and assessments (MEGA, 2014). Task 1 was designed to be the one task not scored by ETS or raters outside of the teacher candidate's university, since it was formative and not definitive. However, the next three tasks, 2-4, were all summative in nature and when MOPTA was implemented in the fall of 2015, would be scored by raters outside of the teacher candidate's university (MEGA, 2014). For clarification, all the scores for the tasks in this study were scored by university faculty and supervisors due to the program in its piloted stage as opposed to full implementation.

Task 2.

The second task was entitled "Assessment and Data Collection to Measure and Inform Student Learning" (MODESE, 2013c, para. 1). Students needed to show their "understanding, analysis, and application of assessment and data collection to measure and inform student learning" (MODESE, 2013c, para. 1). This task focused on state standards 1, 2, 3, 7, and 8 (MODESE, 2013a) and similar to Task 1, asked for a variety of

responses on the topic of data and assessment with artifacts uploaded to support the commentary. In completion of this task, elementary teacher candidates were required to focus on the subject of reading. Below outlines the requirements for textbox 1 of Task 2 in order to provide an example of the type of prompts the students would respond to throughout the task:

Textbox 2.1.1: Selecting a Single Assessment: A. How does this assessment align with standards, learning goal(s), the lesson you are teaching, and student needs? B. What data did you use to establish a baseline for student growth related to this lesson's learning goal(s)? C. Describe the rubric/scoring guide you have selected/designed. How will you communicate its use to your students? D. What evidence of student learning will you collect from this assessment? How will you collect that data? Provide a rationale for your data-collection method. (MODESE, 2013c, para. 3)

In opposition to Task 1, the teacher candidate created all the artifacts for Task 2; there were no templates exemplars to follow. The artifacts required were:

a baseline data document (maximum of two pages); the selected assessment (maximum of two pages); a representative page of the rubric/scoring guide (maximum of one page); a representative page reflecting a graphic representation (e.g., spreadsheet, pie chart, table) of the collected data (maximum of two pages); a student work sample from Focus Student 1 (maximum of one page); and a student work sample from Focus Student 2 (maximum of one page). (MODESE, 2013c, para. 2)

Also in opposition to Task 1 was the summative nature of this task as opposed to the formative nature in Task 2 (MEGA, 2014). The writing and artifacts submitted for Task 2 must exhibit the teacher candidate's process for using data to influence a summative assessment by using baseline data from a previous formative assessment in order to create the artifact needed not only for Task 2 but for effective learning to take place (Margolis & Doring, 2013; Peck & McDonald, 2013).

Task 3.

The third task was similar to Task 2 in that it was also formative, but this addressed "Designing Instruction for Student Learning" (MODESE, 2013d, para. 1). It asked teacher candidates to show their "ability to develop instruction, including the use of technology, to facilitate student learning" (MODESE, 2013d, para. 1). The task also required the elementary teacher candidate to focus on the subject of math. This addressed all of Missouri's teacher standards except for standard 9 (MODESE, 2013a). Below outlines the requirements for textbox 1 of Task 3 in order to provide an example of the type of prompts the students would respond to throughout the task:

Textbox 3.1.1: Standards and Learning Goals and Student Background

Information: A. What learning theory/method will guide your planning process? Provide a brief description of the theory/method. How will you make use of it? B. What learning goal(s) and standards, both Missouri and national, did you identify for the lesson (provide the number and title of each standard that you list)? How will they guide the planned learning activities? C. What is the content focus of the lesson? What related content that the students have previously encountered will support the learning in this lesson? D. What are some difficulties students might

encounter with the content? How do you plan to address these difficulties?

(MODESE, 2013d, para. 3)

The task requirements included the following seven artifacts:

representative pages of your lesson plan (a sample template is provided, but candidates may submit a plan of their own) (maximum of two pages); a plan for differentiation for Focus Student 1 (maximum of one page); a plan for differentiation for Focus Student 2 (maximum of one page); a teacher instructional artifact (maximum of one page); a work sample from a student other than the two Focus Students (maximum of one page); a work sample from Focus Student 1 (maximum of one page); and a work sample from Focus Student 2 (maximum of one page). (MODESE, 2013d, para. 2)

As seen through the task's objective and the artifacts, this focused on effective lesson planning through differentiation. Similar to Task 2, candidates were required to submit work from two focus students. This type of evidence can not only lead the teacher candidate to insights on particular students learning patterns, but showed the raters how effective this teacher candidate would be in his or her own classroom in the future (MEGA, 2014). As seen in textbox 3.1.4, which asked what types of technology would be used and how this would enhance the learning process, technology was emphasized in this task. Researchers have shown the impact technology can have on students in the 21st century (CCSS Initiative, 2014; Margolis & Doring, 2013) and will be expanded upon later in this review of literature. It was also important to note that each task was used as a scaffold from the previous task(s). For example, although Task 3 focused on technology, technology was mentioned in at least one textbox of each task (MODESE, 2013b, 2013c,

2013d, 2013e). Furthermore, the charts used as artifacts from Task 1, were consistently referred to in at least one textbox within each task (MODESE, 2013b, 2013c, 2013d, 2013e). This type of scaffolding created an assessment of connections, as opposed to random questions for each task that were unconnected to the previous task. It also put an emphasis on the formative assessment of Task 1 as it was used to create a starting point for all lessons and activities planned by the teacher candidate during the student teaching practicum (MEGA, 2014).

Task 4.

The fourth and final task focused on “Implementing and Analyzing Instruction to Promote Student Learning” and provided the following instructions to the teacher candidate:

In this task you will demonstrate your ability to plan and implement a lesson using research-based instruction. You will also show how you are able to adjust instruction for the whole class as well as for individual students within the class. Finally, you will demonstrate an understanding of reflective practice. (MODESE, 2013e, para. 1)

Similar to Task 2, this task stemmed from all of the state standards except for standard 9. Just as Task 2 asked the teacher candidate to use data in planning assessment, this task asked the teacher candidate to use research in planning lessons and using effective instructional strategies (MODESE, 2013c, 2013e). Below outlines the requirements for textbox 2 of Task 4 in order to provide an example of the type of prompts the students would respond to throughout the task:

Textbox 4.1.2: Instructional Strategies: A. How do you plan to use academic content language to promote student learning? Provide a rationale. B. How do you plan to engage students in critical thinking to promote student learning? Provide a rationale. C. How do you plan to use questioning skills to promote student learning? Provide a rationale. D. How do you plan to integrate reading into the content you will teach? Provide a rationale. (MODESE, 2013e, para. 4)

In order to show evidence of the teacher candidate's commentary, the following artifacts were required:

one fifteen-minute video (mandatory), which may be a full fifteen minutes (unedited) or may be separated into three five-minute segments (each unedited) combined into one file; two representative pages of your standards-based lesson plan (maximum of two pages) (a sample template is provided, but candidates may submit a form of their own); a student work sample from Focus Student 1 (maximum of one page); and a student work sample from Focus Student 2 (maximum of one page). (MODESE, 2013e, para. 2)

The main artifact, divergent from the previous tasks, was the video submission. However, due to technology policies in some host schools for teacher candidates, MEGA established a Task 4 alternate artifact (Hariston, 2014; MODESE, 2015). The alternate artifact solved policy issues, but research has shown that one of the most effective ways for educators to improve is through observing their own instructional methods (Gates, 2013; Tamer, 2014). Furthermore, the textboxes within the task ask the teacher candidate to refer to various points in the video where they exhibited a particular instructional strategy or viewed a learning activity (MODESE, 2013e). Similar to Tasks 2 and 3 a

lesson plan and student work was asked to be uploaded as artifacts, further exhibiting the importance of using such evidence in the teacher candidate's future career in his or her own classroom (MODESE, 2013c, 2013d).

edTPA

EdTPA was created as a national assessment for states to adopt as a method to evaluate teacher candidates using a performance-based system (AACTE, 2013). Different from MoPTA, edTPA was divided into 27 different tests tailored to the subject area; however, similar to MoPTA, edTPA evaluates teacher candidate's "planning, instruction, and student assessment" through commentary and supporting evidence (AACTE, 2013, p. 1). According to the American Association of Colleges for Teacher Education (2013) edTPA created their assessment by aligning with National Standards (InTASC) and Common Core State Standards, and "shares key points of alignment with the Council for Accreditation of Educator Preparation Standards (CAEP)" (p. 1). Just as MoPTA partnered with ETS to create their assessment, edTPA partnered with the testing company Pearson (edTPA, n.d.), which charged \$300 (edTPA, 2014, para. 1). The test was created due to increasing sentiment that all new educators must be ready to "meet the academic needs of all students" (AACTE, 2013, p. 2). Similar to other companies and places of employment, TEPs are a way to have an agreement with the people of the community and state to hold themselves accountable for the services they provide to the teacher candidates they are preparing (AACTE, 2013). EdTPA insisted they were not just measuring "teaching effectiveness" but using the assessment as a means to gather data to help the teacher candidate grow into a capable educator ready to instruct students and help them reach their full academic level (AACTE, 2013, p. 4).

EdTPA created 27 different subject assessments revolving around the following tasks: “Task 1 Planning: Planning for Instruction and Assessment; Task 2 Instruction: Instructing and Engaging Students in Learning; Task 3 Assessment: Assessing Student Learning” (AACTE, 2013, p. 11). These are parallel to MoPTAs tasks that cover knowing students and their community before planning a lesson (MODESE, 2013b), collecting data to “measure student learning (MODESE, 2013c, para. 1), “designing instruction for student learning” (MODESE, 2013d, para. 1), and “implementing and analyzing instruction to promote student learning” (MODESE, 2013e, para. 1). Also similar to MoPTA were the artifacts required with the commentary, such as lesson plans, video recordings, and examples of student work (AACTE, 2013; MODESE, 2013e). However, edTPA did not offer any alternative to the video recording (AACTE, 2013).

A difference between edTPA and MoPTA was the way they were scored. For edTPA, “each task was evaluated with five separate rubrics...[with] 15 different elements of teaching...scored. A candidate receives[d] a total score on the edTPA that could range from 15 to 75 (AACTE, 2013, p. 12). This varies from MoPTA in that scores were given on a range from 1-4 and a rubric was provided for each textbox within the task (MEGA, 2014). The MoPTA rubric is discussed more specifically in Chapter Three. Furthermore, the edTPA, as stated in AACTE’s 2013 report, “650 teachers and teacher educators score the 12,000 plus candidates who participated in the field tests” (p. 1). Those that scored edTPA also went through extensive training, over 20 hours’ worth, and had a wide variety of resources available to them via Pearson (edTPA, n.d.). This varies from MoPTA in that during their pilot only 60 scorers evaluated the assessment; however, the scoring did take place during scoring sessions involving training, which equaled to

approximately 12 hours within one semester. The results of the field testing of the edTPA revealed higher scores for secondary teaching fields than elementary and that overall performance rated planning the task as the highest, “followed by the instruction task, and then the assessment task” (AACTE, 2013, p. 2). Again similar to MoPTA in that using the results from student’s learning to adjust lessons and plan accordingly was one of the more challenging tasks for teacher candidates.

EdTPA showed benefits as a national program in that it adhered to national standards and provided multiple resources via the testing company Pearson (AACTE, 2013; edTPA, n.d.); facts such as these have led to “622 Educator Preparation Programs in 35 states and the district of Columbia participating in edTPA” (AACTE, 2015, para. 1). In 2013, according to AACTE, edTPA had a total of 26 states participating in edTPA by “exploring or trying out” (p. 6) the assessment; this included the state of Missouri. This number increased to 35 in 2015, according to AACTE; furthermore, the number of states not participating in edTPA stayed consistent from 2013 to 2015 with 16 states (AACTE, 2015, para. 1). One change in the states participating was Vermont began to pilot the program, but Missouri dropped out of edTPA to develop their own assessment thus keeping the number at 15 (AACTE, 2015). In opposition, California went from piloting the program in 2013 to full implementation in 2015, thus adding another choice of TPA along with their already establish CalTPA, PACT, and Fresno Assessment for Teachers (State of California, 2015). The reason for such changes in assessment adoption are unclear, but it may be similar to reasons for many opting out of Common Core State Standards (Cassidy, 2015) in order to have more freedom as an independent state. It was unclear why Missouri chose to adopt its own assessment; however, its parent testing

company ETS has developed the MoPTA as a part of its own national assessment known as the Praxis Performance Assessment for Teachers (ETS, n.d.). Therefore, Missouri was not only piloting its own assessment for the states but piloting an ETS nationally created assessment as competition for edTPA.

Competition is certainly important in the market place, but a given for these types of assessments as they open up opportunities to improve upon the challenges edTPA faced during its implementation (The Warner School of Education, 2015). The Warner School of Education (2015) cited work by their professors, Meuwissen and Choppin, that showed two states, Washington and New York, “felt unprepared during the first year of edTPA implementation” (para. 1). These sentiments were also shared by university supervisors and faculty while they implemented MoPTA, as discussed in Chapter Four. They cited that only 47% of New York assessment participants felt they had a “good understanding” of the national assessment and only 65% reported the same understanding in Washington (The Warner School of Education, 2015, para. 3). Many cited the problem with the quick implementation before any knowledge of edTPA was shared with the teacher candidates; however, it was noted that the roll out of the program was more effective in Washington than in New York (The Warner School of Education, 2015). ETS (n.d.) certainly may have seen these types of challenges as a way to develop a more effective product for states looking for an alternative. California may be at an advantage by offering a variety of tests for their teacher candidates as opposed to just one; however, to compare MoPTA and edTPA to their exam, CalTPA is examined in the next section.

California's Teacher Preparation Assessments

Although no studies were conducted on MoPTA, many studies were conducted on the California state assessment, CalTPA (California Teacher Performance Assessment). The tests for each state, Missouri and California differed, but they both were created with the testing company ETS (MEGA, 2014; State of California, 2015) and resembled the format of edTPA (AACTE, 2013). At the time of this writing, a total of 35 states (including Washington, D.C.) required edTPA as a requirement for certification; with the other 16 states choosing a different TPA as an aspect of their certification requirements and/or educator preparation program (AACTE, 2015). The CalTPA, similar to MoPTA, was broken up into four tasks focusing on subject-specific pedagogy, designing instruction, assessing learning, and the culminating teaching experience (CalTPA, 2013). The tasks were similar in nature to MoPTA in that they required teacher candidates to use evaluation of student learning during instruction and use assessments in lesson planning and analyzing student progress; however, the CalTPA also focused on pedagogy and content more so than MoPTA (CalTPA, 2013; MEGA, 2014). The tasks were also scored on a 1-4 scale by various university supervisors and faculty, again similar to MoPTA (CalTPA, 2013). One difference between California's TPA and other states, such as Missouri, was that California offered options to universities on which assessment to use: CalTPA, edTPA, PACT, and Fresno Assessment for Teachers (State of California, 2015).

In Chung's 2008 study of an earlier ETS exam for California, the Performance Assessment for California Teachers (PACT), suggested that when exams like these were implemented in an educated and thoughtful manner, it added value to the education of future teachers. Chung's (2008) study was qualitative in that it questioned two student

teachers, Tracy and Joy, before the PACT and after. Before completing the tasks the PACT required, Chung asked them their “attitude toward the teaching event” (p. 13). Tracy conveyed a confident attitude whereas Joy’s feelings were varied. During the teaching event their emotions were monitored as well and it was clear that Tracy’s attitude had changed. She felt the work to complete the PACT was “time-consuming” and involved a “heavy workload”; however, the work was not challenging. For Joy, her mixed feelings varied from Tracy in that she found the work “challenging” and “rigorous” (p. 13). However, when emotions were put aside and the true goal of the program was examined, the benefits to PACT were clearly seen. Chung asked the students what learning was achieved from the PACT and both teachers ranked “planning an extended learning segment” and “modifying lessons based on assessment of student learning” at the top of the list (p. 13). These findings concurred with Sandholtz and Shea’s (2012) study of the PACT in that the assessment challenged teachers to show evidence in how they analyzed their own teaching methods and made changes to accommodate students’ struggles. In conjunction to Sandholtz and Shea, Chung asked Tracy and Joy what changes they saw in their instructional strategies from completing the PACT. They both saw the shift from student engagement to evidence of student learning (Chung, 2008; Sandholtz & Shea, 2012). This was clearly described as one of the goals of a teacher preparation program: evidence of learning. Teachers were noted throughout the literature as dynamic, inspiring people, but just engaging students was not enough in this challenging, competitive, 21st century world (Chung, 2008; Sandholtz & Shea, 2012). The teacher is required to cite evidence of student learning and accommodations when assessments reflected a lack of learning. This was what TPAs such as PACT and CalTPA

sought to address (Chung, 2008; ETS, 2014; Sandholtz & Shea, 2012). Although not perfect, the use of TPAs sought to increase the student teacher's success when it was time for he or she to control a classroom of his or her own.

Aligning TEPs to Standards

To create accountability and consistency, TEPs across the nation were aligned to either state standards or national standards. As mentioned earlier, much like CCSS, the CCSSO (2011) formed a consortium that consisted of public officials who worked with various education departments in Washington, D.C. and created a set of national standards for teachers. They entitled these standards "Model Core Teaching Standards" or MCTS (CCSSO, 2011). The purpose was to articulate effective teaching and learning within a transformed public education system that empowered every learner to take ownership of their learning. Furthermore, it emphasized the learning of content and application of knowledge and skills to real world problems that valued the differences each learner brought to the learning experience, and leveraged rapidly changing learning environments to maximize learning and engage learners. A transformed public education system required a new vision of teaching (CCSSO, 2011, p. 3). This change was much needed as seen in the stories told in *Lessons of Hope* by Klein (2014), a former chancellor of the New York City Department of Education, where he stated the many teachers who entered the workforce where "not sufficiently skilled, and too often they lacked real expertise in their subject areas" (p. 55). Klein further cited a principal who fired 80% of her teachers who were "not up to snuff" (p. 56). Clearly, a bold change needed to be implemented and MCTS hoped to evolve the teaching profession for the better.

The decision to create national standards was connected to CCSS in that students were required to meet higher standards. This was true even in states that had not adopted CCSS, or might refute CCSS in the future. Every state, school, and community agreed that setting higher standards for students was a necessity in this highly-competitive, technology driven, 21st century new world (CCSS Initiative, 2014). Therefore, not only did students need to be prepared to think critically and creatively, but their teachers needed to be prepared as well (Chung, 2008; CCSSO, 2011; EdTPA, n.d.; Margolis & Doring, 2013; Robinson, 2014; Sandholtz & Shea, 2012). In order to achieve this, national standards were created. One aspect of the standards was a focus on “personalized learning for diverse learners” (CCSSO, 2011, p. 3) similar to Task 3 for MoPTA (MODESE, 2013d). Instructors needed to be aware of their students’ diverse life experience, work or activities where they excelled, concepts they previously learned, and various values from culture and community that were vital to the educational experience (CCSSO, 2011). Another focal point of the standards was “a stronger focus on application of knowledge and skills” (CCSSO, 2011, p. 4). This involved students taking what they learned in the classroom and application to a larger context: the community, state, nation and world. This involved teachers instructing students on topics such as “problem solving, curiosity, creativity, innovation, communication, interpersonal skills, the ability to synthesize across disciplines, global awareness, ethics, and technological expertise” (CCSSO, 2011, p. 4). This aspect of the standards was in complete compliance with the CCSS for students, and just as state education departments reviewed what students needed to know to succeed in this new world, they also needed to review teacher education programs and examine what instructors needed to know to prepare K-12

students (CCSS Initiative, 2014; CCSSO, 2011). A third focus of the national InTASC standards was creating “a collaborative professional culture” (CCSSO, 2011, p. 4). This required the clarification of a much-needed change for teachers in the classroom: working together, not in isolation. Again, what was true for students was the same for teachers; and as collaboration between pupils resulted in academic success, collaboration between instructors improved instructional strategies and application (CCSSO, 2011; Margolis & Doring, 2013; Robinson, 2014). A final aspect of the standards was the implementation of “new leadership roles for teachers and administrators” (CCSSO, 2011, p. 5). Another shift in education was the changing role of the teacher within the school. According to CCSSO (2011) it was imperative for teachers to be the instructor in the classroom and a leader in the building. By developing teacher leaders it raised the expectation that educators get out of the classroom and pursue more professional development opportunities and educational advancements to improve as an educator, and share what they learned with other educators to raise student achievement (CCSSO, 2011). Through a continued learning commitment by teachers, and a collaborative spirit with peers, teachers had an evolved role that expanded outside of the classroom.

An example of these national standards implemented at the university level was seen at the University of Pretoria (Naidoo, 2012). They made many changes to the curriculum within their School of Education, such as involving “community engagement” in order to “strengthen future teachers’ knowledge of diversity, social justice and themselves” (Naidoo, 2012, p. 78). Naidoo (2012) explained other components of their new curriculum included “the teacher and learner dealing with the pedagogical content knowledge in the classroom and “the teacher trying to understand the learners and the

community they come from” (p. 78). This added facets to their program that not only aligned to national standards (CCSSO, 2011), but it was beneficial since the faculty worked together to design the curriculum (Naidoo, 2012). In comparison, MoPTA although following MoSPE, Missouri standards, as opposed to national standards, also followed these same components as Task 1 that addressed knowing students and their community as well as adjusted the lessons and activities to suit the learner’s background (MODESE, 2013a). Through implementation of new curriculums that follow knowing students and their community through active engagement, the researcher believed universities could produce more efficient educators for the future.

Online Feedback and Foliotek

A crucial aspect to all TPAs were the uploading of commentary and artifacts via online portfolio systems and evaluation of such TPAs with feedback via the online system. EdTPA used their own online portfolio system (edTPA, n.d.) and MoPTA used the online portfolio system Foliotek (2014). This system allowed students to submit writing and artifacts by uploading documents such as word documents, pdfs, images, and videos, and also provided a scoring template for the scorer to use when grading. Besides issuing a 1, 2, 3, or 4 score for each textbox, a section for an overall score for the task was given as well as a box for the rater to type in a comment or upload a document with feedback or commentary. In McVey’s (2008) study on feedback via an online environment, she cited feedback as a crucial component of effective learning. Furthermore, scorers or instructor’s feedback in an online environment may not be as useful as they hoped it would be for students (McVey, 2008). When the instructor is not present to explain the comments and criticism, students can be left with unanswered

questions. McVey explained, “When online feedback is not transparent, students often become anxious and lose motivation because they are confused about what and how well they are doing” (p. 40). However, when it is done correctly, as cited in Wichadee’s (2013) study on wiki sites to improve writing ability, when feedback was collaborative and clear in nature, it was effective. Therefore, it was important for the scorers of the MOPTA tasks to use transparent language in Foliotek for students to understand what needed to be fixed in order for them to resubmit and achieve a passing score. Although the final logistics of MOPTA would not be completed until August of 2015 (MEGA, 2014), it was possible that a score of 1 would need further revisions by the teacher candidate. Therefore, it was up to the scorer to give appropriate feedback so the student could adjust the writing and/or artifacts accordingly. This may be an easier feat when the task was scored by the university supervisor, as in the piloted program; however, when ETS took over scoring in the fall of 2015, they needed a clear understanding of the types of feedback most useful for students to adjust their writing in the most effective way (McVey, 2008). One suggestion from McVee’s research claimed using programs like track changes for Microsoft Word gave students more of a “pen and ink” type of feedback that students were more comfortable with (p. 40) as opposed to the summarized commentary that was given in an online environment, such as Foliotek. A solution for the MOPTA scoring would be to upload the student’s MOPTA tasks and artifacts that have track changes implemented from the scorer. This can be uploaded as a document on the scoring template. Then, students would have clear edits and feedback so if changes were needed they were accurate. Although it is unclear if ETS would allow such feedback, the figure below exemplifies how this could be a possibility.

Unit Expectations Scoring				
	1 - Very Weak (1 pt)	2 - Satisfactory (2 pts)	3 - Good (3 pts)	4 - Excellent (4 pts)
Textbox 1.1.1 Community, District, School Contextual Factors that Influence Instruction	evidence	evidence	evidence	
Textbox 1.1.2 Classroom Demographics and Knowledge of Students	evidence	evidence	evidence	
Textbox 1.2.1 Available Resources to Enhance Student Learning	evidence	evidence	evidence	
Textbox 1.2.2 Student Interest Inventory	evidence	evidence	evidence	
Textbox 1.2.3 Communicating with Students and Families	evidence	evidence	evidence	
Textbox 1.2.4 Rules and Procedures	evidence	evidence	evidence	

Unit Scoring				
	1 - Very Weak (1 pt)	2 - Satisfactory (2 pts)	3 - Good (3 pts)	4 - Excellent (4 pts)
Task 1				

OVERALL COMMENTS

Comments

This is a test


 Task 1 Commentary.docx

Figure 1. Screen shot of Jan Student’s score and feedback for Task.

As seen in the above figure, this screen shot shows how a task is scored via Foliotek. In Jan Student’s (invented student for the purposes of testing MOPTA scoring on Foliotek) Task 1, the student received a score of 4 for each textbox and there was a space at the bottom for feedback. However, there is also a document attached, symbolic of how an evaluator might use track changes via Microsoft Word to give additional specific feedback for each textbox and attach it to the score in case improvements were necessary for a score of 2 or better.

Benefits to TPAs

Collaboration.

One benefit to using assessments such as TPAs, as outlined by Robinson (2014) in an article for *Principal*, was the collaboration between the teacher candidate and the principal. Merritt (as cited in Robinson), a principal in Pasadena, Maryland, explained the intensity shift in working with teacher candidates, and the incredible benefit that came with the more challenging teacher education standards. He stated how it was commonplace to show teacher candidates “his school improvement plan and work with them to develop high-impact strategies for target groups such as boys and students from low-income families” (p. 26). The principal stressed the need for alignment with standards and data provided from assessments. Furthermore, Merritt discussed the constant feedback required of him and the cooperating teachers. In order to help teacher candidates write their reflections and rationales required of TPAs, Merritt had many conversations with the student teachers where he tried to encourage them to think about the “why” behind their teaching strategies (Robinson, 2014). However, the success he witnessed with teacher candidates could not have been done without the entire faculty’s commitment to improving teacher candidates and the collaboration that took place between cooperating teachers and the teacher candidates. Robinson stated, “Teacher preparation must be seen as part of the delivery of student achievement. Preparation programs must take the lead in building and sustaining a climate that makes this possible” (p. 26). Therefore, for any school and future teacher to truly benefit from the TPA, the researcher agreed with Robinson (2014) that everyone must be on board. It was noted this type of collaboration may be too much change for some principals. In a study by Goldhaber and Cowan (2014) they cited surveys from principals indicating they hired new teachers based on other forms of documentation other than TPAs. For example,

cooperating teachers' evaluation, experience, credentials, and other examples involving observations; few cited using portfolios, TPAs, or other unobserved materials. Just as the USDOE report from 2003 endorsed looking at more observable items like types of degrees, principals took this recommendation and noted difficulty as they adjusted to another change (Goldhaber & Cowan, 2014).

Critical thinking skills.

The move to teacher performance assessments also provided for the deeper thinking skills an instructor must have for an effective classroom to produce students who critically think for themselves. TPAs allowed the teacher candidates in their evaluation to “contextualize their teaching more deeply in actual classroom life, connecting their pedagogical decision making to student learning in authentic ways such as videotaping and analyzing classroom experiences” (Margolis & Doring, 2013, p. 272). This was in opposition to the traditional way of evaluating teachers through a portfolio assessment. Although the video component was a step in the right direction as far as using reflection and technology, which research has proven to be effective, it does bring up privacy issues (Pullin, 2014). Pullin (2014) cited TPAs needed to “review state law with local legal counsel to determine the provisions applicable to privacy interests and procedures in school” (p. 17). This type of legal issue did impact the MoPTA in that some schools did not allow videotaping in the classroom in their own district's policies, thus a pre-service candidate teaching in a school with this policy could not submit the required artifact for Task 4 (Hariston, 2014; MODESE, 2015). However, MEGA did rectify this with an alternate artifact for the task (MODESE, 2015). Certainly any legal issues were not anticipated during the move from portfolios to TPAs. In fact, the move away from

portfolios was a result of teacher education programs facing higher standards with the move to CCSS or Common Core State Standards (CCSS Initiative, 2014; Margolis & Doring, 2013); as well as keeping reliability with scoring by using outside raters (AACTE, 2013). The outside raters may also cover any legal issues that may arise from biased scoring or any untrained evaluators scoring assessments (Pullin, 2014). Therefore, the TPA hoped to accomplish the same objective that the Smarter Balance Consortium hoped to attain from the implementation of CCSS. “This initiative aims to create evidence of teaching competence for certification, provide data for teacher education program improvement, and facilitate a learning platform for new and practicing teachers” (Margolis & Doring, 2013, p. 273). This type of initiative was similar to CCSS in that they both wanted to provide evidence of learning. For TPA, it was evidence of the pre-service teacher’s readiness for the classroom; for CCSS, it was evidence of the student’s preparedness of 21st century skills (CCSS Initiative, 2014; Margolis & Doring, 2013). For example, CCSS had students explain their reasoning behind problem solving math equations and TPA had students make connections between philosophies presented in their college courses and their own teaching strategies used in the classroom (CCSS Initiative, 2014; Margolis & Doring, 2013). By using common standards and assessments for all students and teachers in the state of Missouri, there was a collective measurement for all school districts.

Providing and gathering data.

Besides encouraging critical thinking, for both teacher candidates and students, the TPA also had the benefit of providing data and requiring teacher candidates to use data in the classroom (Task 2 for MoPTA). Margolis and Doring (2013) cited having real

numbers from real assessments fostered and encouraged optimistic change for the leaders of a school district. When data regarding teacher candidates was analyzed and shared, it led to a successful program with clear areas to improve and meet expectations (Margolis & Doring; Peck & McDonald, 2013). This shifted the teacher candidate, and the rest of the supportive faculty and staff, out of isolation and into the conversation of improving student learning and preparing teachers for the classroom. If schools were moving to increased accountability with CCSS, the researcher believed so should teacher education programs at the university level. Much as the National Governor's Association came together to create CCSS to hold students to a higher standard, the Council of Chief State School Officers (CCSSO, 2011) gathered to create K-12 national standards for teachers to use to guide their curriculum. This was historic as it was the initial implementation of a community of teachers and leaders came together to have accountability for instructor performance in the classroom (Margolis & Doring, 2013). As Green (2014) noted in her book *Building a Better Teacher: How Teaching Works (and How to Teach it to Everyone)*, teaching is a "voodoo" combination of personality and "passion" is a "dangerous notion" (p. 9). By holding true to the cliché "He who can, does. He who cannot, teaches" implants the idea that teaching is for a stereotyped group where performance cannot be measured (Green, 2014, p. 9). This type of thinking leads to teaching that is not able to be evaluated (how would you measure "charisma and passion"), but by using data and performance assessments real evaluations can take place which leads to "building a better teacher" (Green, 2014, p. 9).

Disadvantages to TPAs

Cost.

One of the concerns regarding the use of MOPTA and other assessments similar is the cost of taking the test. According to Cavanagh (2013) testing companies have seen a great demand for their services due to shifts in local or national policies impacting K-12 and higher education. These changes in policy meant big business for companies that provided testing materials, such as the exam or practice tests (Missouri Association of Colleges for Teacher Education [MCATE], 2014). This was not the first time testing companies experienced an increased need for their services; a similar increase was seen with the passing of the No Child Left Behind Act in 2001 (Cavanagh, 2013). States were unprepared to design, distribute tests and testing materials and hired independent testing companies such as ETS to assist in the redesign (Cavanagh, 2013). The state of Missouri changed their testing requirements in the fall of 2014 as well as the total cost of taking the assessments as discussed in Chapter One.

According to MODESE (n.d.b.) the new certification requirements eliminated the use of the previous test, *Praxis II*, and replaced with the MOGEA, MEP, and MOPTA. The MOGEA and MEP equated to a total of \$7; however this did not include the price of the MOPTA tasks or the Missouri Content Assessment which ranged from \$77 to \$125 dollars. Although registration for the tasks as well as final costs at the time of this study were not yet published, estimates included a total price of \$275 with additional monies being spent if a task or tasks were given a score of 0 or 1 (MCATE, 2014; MEGA, 2014; MODESE, n.d.b.). MCATE (2014) cited many problems with the shift to MOPTA assessments including the cost. They stated:

Teacher candidates in Missouri's institutions of higher education should not pay for the privilege of field testing tools that will ultimately bring in millions of

dollars to ETS. Over the next four years ETS stands to make nearly \$5.5 million from Missouri's teacher candidates... That candidates should be required to underwrite the development of this battery of assessment tools is even more problematic because of the financial burden it will create. (para. 2)

Also important to note was the price of college tuition on top of the costs for these tests. Carey (2015) cited in his book *The End of College* that the price of tuition has raised exponentially. The author claimed it has risen 80% from 1995 to 2013 with a cost of \$18,391 of tuition per year (p. 115). This amount of money, on top of other items such as mandated tests like MoPTA, was alarming. In an unpublished response, ETS cited the research basis for the tasks and that all questions regarding MOPTA, including cost, would be addressed in their Technical Manual to be published after an examination of the piloting of the program during the 2014-2015 school year (MCATE, 2014).

Furthermore, other testing companies such as Smarter Balance which designed for CCSS, claimed the use of technology-based exams required more "complex performance tasks" which required more effort by the testing companies and scorers, thus a higher cost (Gewertz, 2013). This addressed a paradox between concerns for rising costs in education and the need for more data and evaluative formative assessments; in order to have the latter, a price, literally, must be paid for those results. Gewertz (2013) in her study of computer-based exams cited that tests served a purpose of planning curriculum accordingly. Santelises, Education Trust's president of K-12 policy and practice, as cited by Gewertz, asked "Are those students going to have access to the kind of experiences and curriculum that prepare them for those kinds of tasks? Are teachers being prepared to do that?" (p. 5). Although referencing Smarter Balance assessments, this applied to

MOPTA as well, in that the assessment should also serve as preparation at the university to certify highly qualified teachers; just as K-12 should be producing highly qualified students. Santelises clarified this point with “We need to stay focused on the teaching and learning” (as cited in Gewertz, 2013, p. 5) as opposed to the cost of tests or the right technology to support any type of assessment. As with any assessment, the focus must be on the preparedness of the student and the reflection this has on the K-12 or higher education environment (Gewertz, 2013).

Workload.

Teacher candidate.

The amount of time it took for a teacher candidate to run a classroom, plan lessons, and teach to high standards was exhausting; and this did not include the hours spent in the university classroom and the requirements that must be completed in order to obtain certification. In the past, the requirements capsulated within a portfolio—a collection of artifacts or evidence of a successful student teaching experience (Margolis & Doring, 2013; Sandholtz & Shea, 2012). However, with the shift to TPAs, the student teacher’s workload increased. According to Margolis and Doring’s 2013 study, student teachers found the TPA as a great assistance to their establishment as an effective teacher, yet they found the workload overwhelming. They lacked enough time in the days, weeks, and months to complete all the requirements of the TPA. Teacher candidates felt this actually took away from their effective teaching methods because they were more occupied with finishing the tasks and other TPA requirements as opposed to working on their instructional methods (Margolis & Doring, 2013). From the study, one teacher candidate reported: “I am a better student, but a worse teacher. The TPA took

time away from my teaching” (p. 277). However, other student teachers saw a clearer connection between the workload for TPA and the student teaching experience. Many cited the videotaping of lessons as a chore and appeared to be just another item to check off the TPA list; it actually served as a “valuable reflective tool” (p. 278). Furthermore, the TPAs insistence upon examining student work led many teacher candidates to go beyond just grading student’s tests and writing, which led to an increased understanding of what students understood and where students struggled (Margolis & Doring, 2013). Although time consuming, through this study, it was clear the hours spent outside of the classroom for meeting TPA requirements, actually benefited the weeks spent in the classroom.

Cooperating teacher.

The cooperating teacher, which was meant to serve as a mentor to the student teacher and essentially co-teach and assist with all aspects of instruction during the internship, required a great deal of work in addition to the important job of fulfilling regular teaching obligations. This role shifted with the adoption of TPAs. Margolis and Doring, in their 2013 study of TPAs, questioned student teachers about their mentor’s ability to assist with TPAs. Due to the lack of communication between the university and/or TPA representatives, the “mentors knew zero” about the new assessment (p. 279). Mentors admitted to this lack of knowledge, yet wished for training regarding the assessment and guidance on what they could do to support the teacher candidate (Margolis & Doring, 2013). However, at the same time, there was little time for a regular education classroom teacher to spend time outside of teaching, to learn about TPA, when simply learning about the objectives of TPA minimally impacted the teacher candidate.

This type of training was described as not adding to teaching K-12 students, and possibly added extra work with no application to the teacher's already full schedule (Margolis & Doring, 2013; Sandholtz & Shea, 2012). Therefore, the new role of the cooperating teacher essentially fulfilled the same duties, but also helped to prepare the teacher candidate for the TPA. Some educators steered away from this role, but others embraced the change and displayed excitement for the challenging journey that faced the student teacher (Margolis & Doring, 2013; Sandholtz & Shea, 2012).

University supervisor.

The role of the university supervisor for the teacher candidate has been one of a mentor and evaluator, but this evolved with the implementation of TEPs. Asplin and Marks (2013) in their study titled "Increasing the Influence of University Supervisors During Student Teaching" examined the relationship between the university supervisor and teacher candidate and how this affected the transmission from the education received at the university level to the K-12 classroom. Researchers argued that TEPs essentially removed all of the evaluation responsibility from the university supervisor and transferred it to the scorers of the TEP, which was often trained educators working outside of the district (Sandholtz & Shea, 2012). Asplin and Marks cited Marks' earlier 2002 work which suggested the role of the university supervisor during the teacher candidate's practicum experience has been discounted and ignored. As opposed to providing perspective and knowledge, the university supervisor was viewed as an assessor instead of a partner to inspire growth in the teacher candidate (Asplin & Marks, 2013; Ongel, Capa, & Vellom, 2002). However, Fernandez and Erbilgin (2009) clarified the opposite viewpoint and noted that university supervisors were crucial in teacher

candidate's applying their college coursework in education to the K-12 classroom. Asplin and Marks furthered this view and cited the university supervisor's role was priceless when there was a positive relationship between the supervisor and candidate, the candidate respected the supervisor's knowledge and experiences, the supervisor was available to the candidate, and there was consistency between the expectations discussed in the college coursework and the potentials of the teacher candidate's practicum (2013).

More specific results of Asplin and Marks' (2013) study provided a suggestion to universities to foster the relationship between the supervisor and candidate. They encouraged the candidate to be enrolled in a course with the university supervisor prior to the student teaching. Student teachers conveyed having a better rapport with the supervisor and saw them as more well-informed and were more prone to go the supervisor with a question or concern (Asplin & Marks, 2013). Taking a course with the university supervisor would also aid in the relationship with the cooperating teacher, according to Asplin and Marks. This conclusion came from the results, which showed a correlation between the university supervisor and the cooperating teacher, i.e. if the student teacher felt the university supervisor was more knowledgeable then they felt the cooperating teacher was less knowledgeable (Asplin & Marks, 2013). Lindley (2009) agreed with this partnership by citing the importance of a mentor to be more than a "cheerleader" to the mentee, but provide more communication about "concerns and issues" (p. 113). To reduce this negative opinion of the cooperating teacher, Asplin and Marks cited those teacher candidates who took a course with their university supervisor had a more positive view of their cooperating teacher. The researchers did not want to draw conclusions from their "broad, general" study to why this was so, however, they did

note having a good scholarly relationship with the university supervisor may be the transfer of an academic respect for the cooperating teachers since in their results they cited having a positive personal relationship with both (Asplin & Marks, 2013, p. 7). Ensuring teacher candidates take a course with the university supervisor included oversight by the university, and required more involvement early on by the university supervisor.

The university supervisor's role also evolved to that of liaison between the implementation of the TEP and the teacher candidate's completion of the TEP. Although there were other oversights that ensured the TEP had been implemented lawfully and appropriately at the university level (Pullin, 2014) it was crucial for the university supervisor to communicate clearly with the teacher candidate in terms of requirements, due dates, and gathering of artifacts. Also, Missouri utilized the university supervisor as the scorer of Task 1 during the implementation of MOPTA in the fall of 2015. Although this score was not submitted as an evaluation of the university program, the teacher candidate was required to receive a score of 2 in order to continue on with completing Tasks 2-4 for the state to score (MEGA, 2014). The scoring of Task 1 by the university supervisor involved training and time to read and score the tasks, however, it was crucial in assisting university supervisors who prepared teacher candidates for the MOPTA (MEGA, 2014) Also of note was the importance of training to score the exam since students would not want to be treated unfairly or scored differently by another university supervisor. In fact, Pullin (2014) noted in her study of legal implications of licensing educators that "the fairness of the scoring system used to judge candidate submissions can all be potentially subject to claims of denials of fair treatment in violation of the

federal or state equal protection or due process clauses” (p. 4). Finally, in Chung’s (2008) study of the California pre-service assessment, she cited one student fully prepared for her TEP due to her professor’s experience with the assessment. This brought the research full circle; if the university supervisor, who was well-trained and versed in the TEP and also the professor of the teacher candidate he or she supervised, the result was a confident and prepared teacher candidate (Asplin & Marks, 2013; Chung, 2008).

Teacher attrition.

There has never been more pressure on universities to produce top educators ready to stay the course than at the time of this writing, made clear by Goldhaber and Cowan (2014) as they discussed the move by policymakers to distribute monies to educational institutions based upon graduates score on the TEPs. This led to the conclusion and divergence between TEPs across universities with Goldhaber and Cowan’s examination of teacher attrition rates in various programs. Their study aimed at possible inconsistencies between the programs and possible answers to the predicament of teachers leaving the profession soon after beginning their career. These researchers examined 20 programs during a 22-year period in the state of Washington (Goldhaber & Cowan, 2014, p. 449), which lacked significant discrepancies between TEPs. This echoed other controversies surrounding the TEPs, such as the lack of true application between the TEP and the classroom (Cochran-Smith et al., 2011). Goldhaber and Cowan’s findings showed three main differences between programs, which may lead to variant amounts of teacher attrition. First, schools recruited different types of students for their TEP which “linked teacher credentials, salary, and demographic characteristics to the likelihood of attrition” (Goldhaber & Cowan, 2014, p. 450). Second, the type of training

received had a direct impact on a teacher's longevity (Goldhaber & Cowan, 2014), which echoed the sentiments of Cochran-Smith et al.'s (2011) findings regarding TEPs that embraced an educator's individualism and belief system. Finally, Goldhaber and Cowan (2014) sought TEPs connection to teacher attrition in the graduates' school of employment. If the TEP focused on a type of school in the surrounding college's community, the teacher graduate most likely would be unprepared for different environments unless the TEPs made diversity an essential aspect of its training. By creating consistency across these three factors in TEPs, it may lead to a decrease in the turnover rate. In Goldhaber and Cowan's study 15.5% of educators departed their present employment every school year, with 7% leaving the state public school system, which was consistent with national rates (p. 452). However, two of the TEPs examined kept attrition rates at 6.9% and 7.5% respectively (p. 452). The researcher concluded a consistent examination of TEPs that continued to produce long-lasting educators, other TEPs with higher attrition rates could change their practices to lower the number of educators leaving the field each year.

Suggestions for Teacher Recruitment and Retention

In an effort to combat attrition, many researchers and educators have given suggestions for recruiting and keeping soon to be educators and current teachers in the field (Chung, 2008; Cochran-Smith et al., 2011; Darling-Hammond, 2014; Dynarksi, 2014; Sandholtz & Shea, 2012; Zeichner, 2003). The program Teach for America (TFA), which recruited top college graduates and trained them to teach in impoverished and/or underprivileged schools, gave universities advice on creating durable TEPs. Darling-Hammond (2011) contended that TEPs needed to build off of what has worked for TFA

in order to produce a skilled workforce for each municipal in the 21st century. First, TEPs should offer scholarships for teachers who offered their services in high-need areas (Darling-Hammond, 2011). This would solve the problem of having qualified teachers in classrooms where often substitutes are used and create a financial lure for college students pursuing education. Second, there should be more assertive recruitment at the secondary level to find future educators (Darling-Hammond, 2011). Cochran-Smith et al. (2011) in their research of many studies throughout the early 21st century confirmed this as well citing that “the recruitment of appropriate teacher candidates” led to higher retention rates (p. 24). The American Association of Colleges for Teacher Education (2013) agreed with Darling-Hammond (2011) and Cochran Smith et al. in their mission to “recruit and retain diverse candidates” as a must to produce effective teachers for all types of students in a variety of school districts (p. 4). College scouts have consistently gone into high schools to find the next great basketball or football player, why not use this same model for the next great educator? Third, there should be more collaboration between universities and current educational institutions, such as hospitals for pre-med students, in order to create a more “practice-based coursework” for education students (Darling-Hammond, 2011, p. 26). The researcher believes this could create a partnership between university and public schools to work together consistently in order to produce a qualified workforce. Fourth, Darling-Hammond cited teacher performance assessments being used nationwide. Note this was from 2011 during the time when only half the nation used these types of assessments, and even cites California’s assessments as a piloted program 20 others states were using as a model (Darling-Hammond, 2011; Sandholtz & Shea, 2012). Fifth, Darling-Hammond discussed holding universities

accountable for the results of accreditation using assessments and other evidence-based approaches to show teacher effectiveness and preparedness. Finally, a fair salary and top working environments were noted for education students to not only pursue the profession but to remain teachers for a quality length of time (Darling-Hammond, 2011).

However, Cochran-Smith et al. (2011) discussed the controversy behind laying out such suggestions like Darling Hammond (2011) taking into consideration that TEPs may not be useful at all due to a lack of evidence of any connection to student achievement. Many of the changes in education at the beginning of the 21st century involved changing from training “highly qualified teachers” to training “highly-effective teachers” (Cochran-Smith et al., 2011, p. 19). Further, the change from a focus on actual teaching tasks to theories and strategies made many education leaders weary of the future (Cochran-Smith et al., 2011). This could be discounted if expert educators left college and entered the classroom ready for a long career, but this was not the case; as of 2002, 46% of new teachers left the profession within five years (Cochran-Smith et al., 2011, p. 19). It was clear that one of the goals of TEPs must be combating teacher attrition as discussed in the previous section.

Due to these concerns, Cochran-Smith et al. (2011) completed a comprehensive examination of research regarding the effectiveness of TEPs with the results published in an article entitled “Teacher’s Education, Teaching Practice, and Retention: A Cross-Genre Review of Recent Research.” One of the aspects of TEPs focused on the graduates of the programs. The research of Hoffman et al. showed that teachers who implemented the practices from their TEP into the classroom were more effective than those who did not use the practices (as cited in Cochran-Smith et al., 2011). Furthermore, those who

were most successful were the ones who contributed their university and TEP with supporting their own philosophies and strategies (Cochran-Smith et al., 2011). In juxtaposition, Steele's study clarified that when teachers left their TEP education behind once they entered the classroom, it was due to the school's lack of support for their beliefs (as cited in Cochran-Smith et al., 2011). When a teacher was comfortable with his or her own belief system, he or she adapted to different learners and behavioral problems; however, when the belief system was not established or fostered, the diversity of learners became overwhelming and leaving a teacher's education program failing the student's in the long run. "As teachers became more comfortable in their schools, they began more meaningful implementation of the practices learned during preservice preparation" (Cochran-Smith et al., 2011, p. 25). Therefore, the researcher concluded that students must not only have a TEP that prepares them for the classroom, but they must have a TEP that they can use as a guideline for the rest of their career.

Summary

In order for students to become successful citizens of the future, they must have highly qualified teachers in the classroom. One way to produce this type of valued instructor is for states to produce and implement highly effective teacher education programs. It was clear from the research that these types of programs were created in various states and produced a more qualified teacher for the student of tomorrow (Chung, 2008). As the country moved towards Common Core State Standards for K-12 students and Model Core State Standards for university students, it was necessary to provide an assessment tool that verified these standards were implemented correctly, appropriately, and efficiently. States such as California implemented the California Teacher

Performance Assessment, while others adopted the nationwide test, edTPA (Chung, 2008; edTPA, n.d.; Sandholtz & Shea, 2012). Missouri adopted its own assessment via a partnership with ETS called the MoPTA and it began piloted implementations around the state 2014. This study aimed to investigate the pilot of this implementation at one private university in Missouri. The next chapter outlined the methodology used for this study.

Chapter Three: Methodology

Purpose

The purpose of this mixed-methods study was to investigate the effectiveness of a private Midwestern university's teacher preparation program through examining the results of the piloted MoPTA during the fall of 2014 and spring of 2015. As stated by Creswell and Plano Clark (2011), a mixed-methods study enables the researcher to use both quantitative and qualitative data in one study in order to produce various types of results so analysis can further the understanding. The quantitative aspect of this study enabled the researcher to examine scores of student teachers to see which tasks fared more difficult than others in order to address possible gaps in the current courses offered to education students seeking certification. Also, this type of study allowed the researcher to compare the scores of an undergraduate student versus a graduate student, to see if the experience of graduate students offered an advantage over undergraduate students in terms of readiness for would-be instructors. Furthermore, the quantitative study led to inter-rater results in order to examine the similarity or lack thereof between university supervisors and faculty scores of the same student. The data collected could lead to possible changes in the teacher education program and training for the university supervisors of these student teachers. The qualitative aspect of this study produced feedback from the supervisors of student teachers regarding the teacher education program, level of preparedness for student teachers, the MoPTA itself, and using the online portfolio system Foliotek. Naidoo (2012) in the study of curriculum change and faculty members as the guiding force, claimed qualitative data such as constructive criticism can "open the doors for communication and thereby improve the chances of the

adoption of change” (p. 71). The researcher hoped through the feedback of university supervisors and faculty, changes needed in the teacher education program could occur in a constructive and efficient matter.

Currently, there were no studies on MoPTA so this study aimed to address the aspects of the assessment and to investigate a university’s preparedness for the assessment, which began implementation in the fall of 2015. By examining other studies on TPAs such as in California (Chung, 2008; Sandholtz & Shea, 2012) the researcher was able to learn the process of evaluating an assessment. Through this examination, universities and school districts could learn about MoPTA, choose to implement curriculum changes that offered support to student teachers who completed the assessment tasks to their highest ability.

Surveys

Once the researcher received approval from the Institutional Review Board of the study university, as well as permission to use the university as a study site (see Appendix B), university faculty and supervisors were asked to answer a voluntary four-question survey (see Appendix A). The researcher developed the survey from the original survey designed by the Assistant Dean of Education. If they did complete the survey, an informed consent (see Appendix E) to use the content of the survey was completed and returned to the researcher. The researcher expected a minimum of 40 completed surveys each semester, however approximately 20 were received. The survey was completed after each scoring session (session 1: Task 1; session 2: Tasks 2-3; session 4: Task 4), during the fall and spring semesters. The researcher also took observational notes (see Appendix

C) during the scoring sessions held at the private Midwestern university. However, there was no qualitative data recorded for Task 4 in the spring.

Scores

University supervisors and faculty scored approximately 276 students' Tasks 1-4 of the MoPTA pilot. Approximately 132 students took the pilot in the fall and approximately 144 students took the pilot in the spring. All students were required to complete Task 1, but were then assigned one other task to complete (Task 2, 3, or 4). Task 4 required a video to be uploaded as an artifact. When MoPTA was implemented in the fall of 2015, students were required to submit all four tasks; however, due to technology policies in some host schools for teacher candidates, MEGA established a Task 4 alternate artifact (Hariston, 2014; MODESE, 2015).

Methodology

The scores from the tasks were received through the online portfolio system, FolioTek. Teacher candidates submitted their responses to the tasks and accompanying artifacts (Word documents and/or PDFs) through FolioTek and university supervisors and faculty scored each task through FolioTek as well. Once all scores were submitted, the Assistant Dean of Accreditation and Faculty Development requested the data from FolioTek; once data was sent to the Assistant Dean, it was scrubbed for anonymity and coded so the researcher could analyze all data. The Assistant Dean did not exclude any data; only student names were removed to protect their identity.

In order to train university supervisors/faculty in scoring the MoPTA, examples were given during each scoring session by the Assistant Dean, it was then scored individually, and scores were discussed in small groups as well as together. Furthermore,

the scorers examined the below rubric provided by MEGA (2014). Although the figure below only displays the rubric for Task 1, MEGA provided rubrics for each task and each textbox within each task online and they were printed for the university supervisors/faculty during the scoring sessions.

Step 1: Knowledge of Students (textboxes 1.1.1 and 1.1.2)			
Score of 1	Score of 2	Score of 3	Score of 4
A response at the 1 level provides minimal evidence that demonstrates the teacher candidate's ability to identify one instructional strategy and one learning activity for a selected community, district, AND school factor; to identify one instructional strategy and learning activity for a selected classroom factor and knowledge of students factor that would further student learning; and to connect the selected factor to each instructional strategy and learning activity.	A response at the 2 level provides partial evidence that demonstrates the teacher candidate's ability to identify one instructional strategy and one learning activity for a selected community, district, AND school factor; to identify one instructional strategy and learning activity for a selected classroom factor and knowledge of students factor that would further student learning; and to connect the selected factor to each instructional strategy and learning activity.	A response at the 3 level provides effective evidence that demonstrates the teacher candidate's ability to identify one instructional strategy and one learning activity for a selected community, district, AND school factor; to identify one instructional strategy and learning activity for a selected classroom factor and knowledge of students factor that would further student learning; and to connect the selected factor to each instructional strategy and learning activity.	A response at the 4 level provides consistent evidence that demonstrates the teacher candidate's ability to identify one instructional strategy and one learning activity for a selected community, district, AND school factor; to identify one instructional strategy and learning activity for a selected classroom factor and knowledge of students factor that would further student learning; and to connect the selected factor to each instructional strategy and learning activity.

Rubric — Task 1 Knowledge of Students

Figure 2. Task 1 rubric (MEGA, 2014).

Besides copies of the rubric, scorers were also given a sheet to record feedback and scores as they read the tasks and examined the artifacts. This allowed them to keep track of thoughts as they were reading so they could then put their scores and comments on Foliotek once all the textboxes were read. Once the scorers completed the training, they were ready to score the tasks they were assigned. Tasks were assigned by the Assistant Dean of Accreditation and Development so scorers could be given specific students to score. For example, a supervisor of a student teacher would be assigned that student

teacher's task; however, it was anonymous. Furthermore, if during the scoring session, a university supervisor or faculty member found a task to be blank or were not assigned any tasks, the Assistant Dean needed to assign a task immediately.

Once scores were received, the researcher compared each task's scores using an analysis of variance (ANOVA) and used this same method for comparing undergraduate students' scores with graduate students' scores. As stated earlier, inter-rater reliability was also examined using Pearson ρ correlation coefficient. Finally, the researcher used surveys and observational data in regards to scoring the MoPTA and overall comments regarding the preparedness of student teachers via the teacher preparation program at this university.

Hypotheses

Hypothesis 1: There will be a difference of scores between tasks (i.e. Task 1 to Task 2).

Null Hypothesis 1: There will be no difference of scores between tasks (i.e. Task 1 to Task 2).

Hypothesis 2: There will be a difference in MoPTA scores between undergraduate teacher candidates and graduate teacher candidates.

Null Hypothesis 2: There will be no difference in MoPTA scores between undergraduate teacher candidates and graduate teacher candidates.

Hypothesis 3: There will be a difference in MoPTA scores given for the same teacher candidate by university supervisors and faculty (i.e.: student A is given the same score by two different raters [university supervisor and/or faculty]).

Null Hypothesis 3: There will be no difference in MoPTA scores given for the same teacher candidate by university supervisors and faculty (i.e.: student A is given the same score by two different raters [university supervisor and/or faculty]).

Hypothesis 4: There will be no difference in MoPTA scores between K-12, early childhood and elementary, middle, and secondary teacher candidates.

Null Hypothesis 4: There will be no difference in MoPTA scores between K-12, early childhood and elementary, middle, and secondary teacher candidates.

Research Questions

Research Question 1: How do university supervisors perceive the process of evaluating teacher candidates' MoPTA tasks through the online portfolio system Foliotek?

Research Question 2: How did faculty change the content of their lessons after evaluating teacher candidates' completed task(s)?

Research Question 3: How do faculty and university supervisors perceive the teacher candidate preparation process (at this particular university)?

Research Question 4: After participating in the piloted program with accompanying training, how do faculty and university supervisors perceive teacher candidates' level of preparation for full MoPTA implementation in the fall of 2015?

Research Question 5: After participating in the piloted program with accompanying training, how do faculty and university supervisors perceive their preparation for full MoPTA implementation in the fall of 2015?

Limitations

Inter-rater reliability was based on one student's scores from multiple raters, however, there were a few students who were only scored by one rater and due to the randomness of assigning the same task to multiple raters once or twice the same rater was given the same student's task to score twice. Although this happened a limited amount of times, it may impact a true inter-rater reliability. Also, sometimes a faculty member did not have time to score all assigned tasks, so a task that was to be scored by multiple raters might have only been scored once. It is also important to note the amount of students given to each university supervisor and/or faculty member to score during the fall session. For Task 1, there were 58 scorers for 132 students; some scorers were given as few as two to score, whereas one was given 16 to score. This did not happen again, as Task 2 yielded 46 scorers for 51 students, Task 3 had 46 scorers for 53 students, Task 4 noted 37 scorers for 30 students; and for Tasks 2-4 some scorers were only given one student to score with a maximum of five students to score. Therefore, Task 1 for the fall may have produced invalid inter-rater results; however, when applied to the scores overall, the difference was minimal. Also, due to the amount of tasks given to each scorer to evaluate, the spring semester produced few multiple raters for one student's task; therefore, the researcher only measured inter-rater reliability in the fall and not the spring. Furthermore, in regards to Null Hypothesis 2, for the fall MoPTA, students were required to declare the degree they were seeking (BA or MA) in Foliotek; however, due to a lack of disclosure of degree seeking in Foliotek for the spring MoPTA, data was only compiled for the fall. Similar in nature, for Null Hypothesis 4, the data was only able to be compiled by elementary and secondary/K-12 for the spring due to students' disclosure

of their certification in the spring, but not required for the fall. Finally, some tasks had an even number of textboxes, so if the score for each textbox ranged from 2-3 then the scorer had to gauge what score to give the task overall, a 2 or a 3. This would have been based on best judgment as opposed to calculated mean.

All surveys distributed to the university supervisors and/or faculty who participated in the scoring sessions were not returned. Out of the 60 surveys distributed during each scoring session, approximately 20 were returned. However, those that were returned included lengthy feedback and this along with the PI's observational notes were enough to justify attributable qualitative data to the study. Also, there was not a Task 4 scoring session in the spring, due to availability of scorers and an understanding by the scorers with evaluating the task, therefore no surveys were distributed or qualitative data collected during that session.

Furthermore, the responses to the survey questions may have been impacted by various technology malfunctions during the start of each scoring session. These are outlined more clearly in the results section of the dissertation, but it was difficult for some supervisors to log in to the computers and access Foliotek with the appropriate username and password. Technology issues with specific computers in the labs also proved a challenge. Due to these delays, some surveys were not completed and some reflected negatively on the portfolio system Foliotek because of these delays as opposed to assessing the system as a tool for MOPTA.

Finally, ETS made slight changes to MoPTA during the piloted program at the study university. Therefore, one version of the assessment was given in the fall and the second given in the spring. However, each task assessed the same area as well as each

textbox; it was a matter of changing the individual textbox prompts to better address each task's objective. For example, Task 2.1 in the spring required inclusion and reflection on baseline data, but the fall Task 2.1 did not.

Teacher Preparation Assessment's Reliability and Validity

When examining TPAs, it was imperative to examine the challenges and validity of the implementation and analysis of results these assessments produced. The study conducted in California by researchers Riggs, Verdi, and Arlin (2009) investigated the "reliability, validity, and procedural adequacy of the teacher performance assessment exam" (p. 13). Riggs et al. used extensive quantitative statistics in calculating the reliability and validity of the test scores as well as the inter-rater reliability between scorers of the exam.

Inter-rater reliability.

The inter-rater reliability was calculated by having the CalTPA coordinator select five tests graded by a different university supervisor or faculty member and then blind scored them her/himself (Riggs et al., 2009). In Riggs et al. (2009) this resulted in 80 tests being scored by the TPA coordinator (p. 23). The scoring of each task resulted in a 1 (lowest) to a 4 (highest), this was the same on the CalTPA and the MoPTA. The table below reflected the inter-rater reliability coefficient for each task's global score, mean score, as well as the specific indicators for Tasks 3 and 4. For purposed expressed by Riggs et al., there were no inter-rater results for Task 2 due to a low sample size. To calculate the reliability, a Pearson correlation was computed. "Intra-class correlations (ICCs) and Pearson correlations were computed between the coordinator's score and the

scores of the original scorers” (Riggs et al. 2009, p. 23). Other researchers, such as Pullin (2014) have run similar inter-rater reliability analyses.

Table 3.

CalTPA Inter-Rater Reliability Coefficient

Score	ICC	Pearson's r
Global Score Task 1	0.25	0.41
Mean Score Task 1	0.28	0.66
Global Score Task 3	0.41	0.61
Mean Score Task 3	0.23	0.43
Goal Setting Task 3	0.18	0.35
Planning for Assessment Task 3	0.04	0.07
Learning About Students Task 3	0.27	0.44
Making Adaptions Task 3	0.21	0.40
Analyzing Student Evidence & Assessment Task 3	0.37	0.48
Reflection Task 3	0.32	0.38
Global Score Task 3	0.27	0.36
Mean Score Task 4	0.32	0.42
Goal Setting Task 4	0.08	0.13
Learning About Students Task 4	0.31	0.37
Classroom Environment Task 4	0.10	0.10
Planning for Instruction Task 4	0.41	0.47
Making Adaptations Task 4	0.19	0.24
Pedagogical Skill Task 4	0.01	0.02
Analyzing Evidence of Student Learning Task 4	0.33	0.39
Reflection Task 4	0.45	0.51

Note. Riggs et al. (2009, p. 23).

The Pearson correlation indicated the amount of variability in scoring for each task. For Riggs et al. (2009) to calculate the Pearson coefficient, Cronbach's alpha (as cited by Riggs et al., 2009) was used with the following criteria: “below .40 is poor, .40 to .59 is fair, .60 to .74 is good, and .75 and above is excellent” (pp. 23-24). For there to be a high correlation between scorers, or a similarity in the scores given by different raters for the same student, the r score should fall between .5 and 1; a medium correlation falls between .3 and .5; whereas a low correlation is below a .3 with a 0 signifying no correlation (Bluman, 2013, pp. 531-539). It is important to note most students were given multiple raters in the fall of 2014; however, due to lack of interest in scoring more

assessments, the Assistant Dean of Accreditation and Faculty Development assigned assessments to fewer raters so there was a lack of multiple scorers for one student.

Therefore, data for inter-rater reliability was only compiled for the fall. Results for inter-rater reliability are detailed in Chapter Four.

Purpose.

Furthermore, Sandholtz and Shea (2012) addressed validity concerns in terms of the concept behind the purpose of TPAs. The results of their study lacked a reasonable and measurable way to examine whether or not teachers were able to implement the teaching strategies and the accommodations proposed in their TPA. The teacher candidates simply stated what they would do in situations where students struggled, but there was no actual implementation of this accommodation and therefore no data to show whether or not it was effective (Sandholtz & Shea, 2012). Riggs et al. (2009) found similar results in their study. Although they focused on the validity of scores, their concerns mirrored that of Sandholtz and Shea. Their study's conclusions recommended others to be wary of the global scores taken from the TPA because they "only provide evidence that the student is failing, but not why or how" (Riggs et al., 2009, p. 35). This coincided with Sandholtz and Shea since both groups of researchers concurred the validity of the scores was contingent upon exactly what the numbers expressed. At the time of this writing, the TPA exams lacked specific feedback to the teacher candidate leading to a possible improvement (Sandholtz & Shea, 2012). Similar to Sandholtz and Shea's concern of not knowing whether teachers implementation of accommodations worked, Riggs et al. expressed the TPA results did not show whether the teacher candidate was ready for the classroom; and if the teacher candidate was not ready, there

was no feedback or conclusions to be drawn from the TPA to guide them in the right direction.

The Research Site and Participants

The researcher observed the piloted program at the study university located in the Midwest. Its Educator Preparation Program is accredited by the Commission on Institutions of Higher Education of the North Central Association Colleges and MODESE, as well as a member of TEAC (Private University, n.d.b., para. 1). Students who took the MoPTA were enrolled in the School of Education. In order to take the MoPTA students had to complete “Stage One” of the Teacher Education Program. It consisted of the following requirements to be met before moving onto “Stage Two” in which the MoPTA would be taken. Students needed to complete a “majority of general education requirements” and complete 12-15 hours “of coursework from the School of Education”; needed a cumulative GPA of 2.75 and content area GPA of 3.0; passed criminal background clearance; taken the MEP; and has taken the MODESE required assessments (Private University, n.d.b., paras. 5-9). Once these are met, students move onto Stage Two where the MoCA and MoPTA are taken (Private University, n.d.b., para. 10).

Students were able to complete MoPTA tasks on any computer, at home or on campus. Once students completed the first MoPTA task, approximately 60 university supervisors/faculty were assigned student(s) to score their Task 1. The university supervisors and faculty who scored the MoPTA consisted of members of the university who were either employed full or part-time as faculty or adjuncts. Many university supervisors and faculty that participated in scoring the MoPTA supervised student

teachers, and many taught education courses; although some of the adjuncts were retired educators who only supervised student teachers instead of teaching as well, and some faculty only taught education courses and did not supervise student teachers. University supervisors and faculty were asked to attend a Task 1 scoring session on September 5, 2014 to look at an example, score it together, and go through the scoring process using the online system Foliotek. It was important to note that the submitting of tasks via Foliotek was deadline driven, meaning if it was submitted one minute after the deadline the entry was blank. The scoring session took place in a room located on campus, which had computers for all the scorers. For convenience, there were three scoring sessions offered throughout the day. The researcher provided the examples, and the scoring sessions were led by the Assistant Dean of Accreditation and Faculty Development. This same format occurred for the Tasks 2-3 scoring session on October 10, 2014, and the Task 4 scoring session on November 14, 2014; and in the spring the Task 1 session occurred on February 6, 2015 and Tasks 2-3 scoring session occurred on March 13, 2015. However, due to availability and an understanding of how to evaluate the task, a spring Task 4 scoring session did not take place. Scorers were given instructions and examples via email/Blackboard and scored them at home. Therefore, no qualitative data was recorded from Task 4 during the spring of 2015.

Summary

MoPTA was piloted at a private Midwestern university in the fall of 2014 and spring of 2015. The researcher used this piloted program to investigate the effectiveness of teacher preparation programs, consistency of scores within multiple raters for one student, and to gain feedback from university supervisors and faculty regarding scoring

the MoPTA and the university's preparedness for this assessment. A mixed-methods approach was used to gain test scores as well as feedback. This type of method allowed the pilot to be examined not only through scores, but also by gaining on site feedback during the scoring sessions. The next chapter explained the results attained from this mixed-methods study.

Chapter Four: Results

Overview

Teacher candidates submitted their tasks through Foliotek, and university supervisors and faculty scored each task through Foliotek as well. Once all scores were submitted, the Assistant Dean of Accreditation and Faculty Development at the researched university requested the data from Foliotek; where all identifiers were removed so the researcher could analyze the data while protecting participants' anonymity. Voluntary surveys were distributed to university supervisors and faculty after each scoring session and were returned to the researcher before leaving, or a few emailed the researcher responses within a week of the scoring session. The researcher also took observational notes during the scoring sessions.

Null Hypothesis 1

The student scores from each task were analyzed to investigate if there was a difference between the scores of each task, overall.

Hypothesis 1: There will be a difference of scores between tasks (i.e. Task 1 to Task 2).

Null Hypothesis 1: There will be no difference of scores between tasks (i.e. Task 1 to Task 2).

As stated in Chapter Two, each task examined a different component of teaching. Task 1, summative in nature, focused on the teacher candidates "knowledge of the students and their learning environment" (MODESE, 2013b, para. 1); Task 2, formative in nature, focused on "assessment and data collection" (MODESE, 2013c, para. 1); Task 3, formative in nature, focused on "designing instruction for student learning"

(MODESE, 2013d, para. 1); and Task 4, formative in nature, focused on “implementing and analyzing instruction” (MODESE, 2013e, para. 1). The highest score a student could achieve was a 4, with the lowest a 1. Individual student scores for each task were entered into the ANOVA calculator, and the following table displayed the results for students completing their student teaching in the fall of 2014.

Table 4.

Results of Student Scores from Tasks 1-4: Fall 2014

Groups	Count	Sum	Mean	Variance
Task 1	241	616	2.56	0.7812
Task 2	90	215	2.39	1.0268
Task 3	87	233	2.68	0.8952
Task 4	42	126	3	0.4878

Note. Task 1: $n=132$; Task 2: $n=51$; Task 3: $n=53$; Task 4: $n=30$. Since one student’s task was scored by multiple raters, the “count” in the table is higher than the population (n).

The results of Table 4 show the number of students who took the task (count), the total of the scores (sum), the average of the scores for each task (mean), and the amount of difference between the scores of each task (variance). The exact same ANOVA test was run for students completing their student teaching in the spring of 2014. The results are displayed below in Table 5.

Table 5.

Results of Scores from Tasks 1-4: Spring 2015

Groups	Count	Sum	Mean	Variance
Task 1	185	479	2.59	0.8847
Task 2	73	185	2.53	0.8634
Task 3	69	161	2.33	0.9902
Task 4	37	111	3	0.6667

Note. Task 1: $n=142$; Task 2: $n=53$; Task 3: $n=49$; Task 4: $n=35$. In some instances one student’s task was scored by multiple raters, thus the “count” in the table is higher than the population (n).

The results of Table 5 show the number of students who took the task (count), the total of the scores (sum), the average of the scores for each task (mean), and the amount of

difference between the scores of each task (variance). Results from Tables 4 and 5 revealed an observable difference between the number of students who submitted Task 1 and all other tasks. All teacher candidates were required to submit Task 1 and then were assigned either Task 2, 3, or 4 as a part of the pilot. An observable examination of these numbers, revealed a small difference; however for a more specific analysis an ANOVA test was completed. Table 6 displays results from tasks completed by students completing their student teaching in the fall of 2014. “Groups” signifies “tasks,” thus a difference between groups is synonymous with a difference between tasks.

Table 6.

Results from ANOVA Test for Tasks 1-4: Fall 2014

Source of Variation	SS	df	MS	F	P-value	F _{crit}
Between Groups	11.650569	3	3.8835	3.255	0.0220	2.633
Within Groups	375.87117	315	1.19324			
Total	387.52174	318				

The exact same ANOVA test was run for students completing their student teaching in the spring of 2015. Just as in the fall, “groups” signifies “tasks,” thus a difference between groups is synonymous with a difference between tasks.

Table 7.

Results from ANOVA Test for Tasks 1-4: Spring 2015

Source of Variation	SS	df	MS	F	P-value	F _{crit}
Between Groups	10.86676	3	3.6623	3.150	0.0255	2.637
Within Groups	316.2761	275	1.15009			
Total	327.1429	278				

Tables 6 and 7 listed the ANOVA results from the Tasks 1-4 scores. Since the F value was less than the F critical value, Null Hypothesis 1 must be rejected. This was concurred by looking at the p-value. The p-value is used to test the strength of the evidence and

works between a range of 1 and 0. A value equal to or less than .05 shows strong evidence against the null hypothesis. Therefore, in Tables 6 and 7 the p-value was less than .05, and solidified the rejection of Null Hypothesis 1.

Tables 8 and 9 closely examined the difference of means between each task using the Scheffe test, a test used with ANOVA when different sample sizes were used. In the fall, there were no differences between the means of the tasks except when examining Task 2 versus Task 4. Task 2 overall had the lowest score as an overall mean, and Task 4 had the highest score as an overall mean. The Scheffe test revealed a significant difference between the scores of these two tasks.

Table 8.

Scheffe Test: Tasks 1-4 Fall 2014

Tasks	F_s	F_{crit}	Significance?
Task 1 vs. Task 2	1.5339105	7.900	No
Task 1 vs. Task 3	0.7992462	7.900	No
Task 1 vs. Task 4	5.9086043	7.900	No
Task 2 vs. Task 3	3.1022223	7.900	No
Task 2 vs. Task 4	8.9625123	7.900	Yes
Task 3 vs. Task 4	2.4588267	7.900	No

By examining the critical value (F_{crit}) it determines the significance level as a limit between the tasks that either show a significant difference or do not. If the calculated value from the test (F_s) is less than the critical value, then the researcher fails to reject the null hypothesis.

Table 9.

Scheffe Test: Tasks 1-4 Spring 2015

Tasks	F_s	F_{crit}	Significance
Task 1 vs. Task 2	0.137391	7.912	No
Task 1 vs. Task 3	2.860514	7.912	No
Task 1 vs. Task 4	4.524499	7.912	No
Task 2 vs. Task 3	1.244993	7.912	No
Task 2 vs. Task 4	4.631379	7.912	No
Task 3 vs. Task 4	9.307408	7.912	Yes

As seen in the above table (Table 9), in the spring, there were no differences between the means of the tasks except when examining Task 3 versus Task 4. Task 3 overall had the lowest score as an overall mean, and Task 4 had the highest score as an overall mean. The Scheffe test revealed a significant difference between the scores of these two tasks. Whereas in the fall, Task 2 showed the highest mean over Task 4. Both semesters showed the highest results for Task 4, but the lowest task varied between each semester. A summary of the results of Null Hypothesis 1 along with recommendations is stated in Chapter Five.

Null Hypothesis 2

This data was analyzed to investigate if there was a difference between the scores of undergraduate students and graduate students, for each task.

Hypothesis 2: There will be a difference in MoPTA scores between undergraduate teacher candidates and graduate teacher candidates.

Null Hypotheses 2: There will be no difference in MoPTA scores between undergraduate teacher candidates and graduate teacher candidates.

The purpose of this test was to see if students at the graduate level performed at a higher level than those at the undergraduate level. The outcomes could reveal if the level of rigor at the graduate level produced a higher score or if the graduate or undergraduate level showed no difference between the scores received on a task. For the fall MoPTA, students were required to declare the degree they were seeking (BA or MA) in Foliotek; however, due to a lack of disclosure of degree seeking in Foliotek for the spring MoPTA, data was only compiled for the fall. Individual scores for undergraduate (Bachelor of Arts [BAT]) students and graduate (Master of Arts [MAT]) students for each task were

entered into the ANOVA calculator, and the following tables displayed the results. Table 10 showed the overall results from the ANOVA test, which displayed the number of students who took each task, by graduate or undergraduate level, as well as the average of the scores.

Table 10.

BAT v. MAT Students

Tasks	Count	Sum	Mean	Variance
Task 1: BA Students	159	393	2.47	0.6938
Task 1: MA Students	77	218	2.83	0.7737
Task 2: BA Students	60	140	2.33	1.0734
Task 2: MA Students	30	75	2.5	0.9482
Task 3: BA Students	58	145	2.5	0.8860
Task 3: MA Students	29	88	3.03	0.7488
Task 4: BA Students	32	95	2.97	0.4829
Task 4: MA Students	10	31	3.1	0.5444

Note. Task 1: $n=132$; Task 2: $n=51$; Task 3: $n=53$; Task 4: $n=30$. In some instances one student's task was scored by multiple raters, thus the "count" in the table is higher than the population (n).

Table 11 closely examined the difference of means between the graduate or undergraduate level of students within each task using the Scheffe test, a test used with ANOVA when different sample sizes were used.

Table 11.

Scheffe Test: BA v. MA Students

Tasks	F_s	F_{crit}	Significant Difference?
Task 1: BA vs. MA	6.965	3.895	Yes
Task 2: BA vs. MA	0.5382	3.949	No
Task 3: BA vs. MA	6.5690	3.953	Yes
Task 4: BA vs. MA	0.2642	4.085	No

There were no differences between the means of the tasks except when examining Task 1 and Task 3. When examining Table 11's overall task scores it showed graduate students scored significantly higher on Task 1 and Task 3. However, there was no significant

difference between Task 2 and Task 4. Therefore, the null hypothesis was rejected for Task 1 and Task 3, but was not rejected for Task 2 and Task 4.

Null Hypothesis 3

To analyze the inter-rater reliability between university supervisors and faculty and their scores for the same student, the researcher used a Pearson ρ correlation coefficient for inter-rater reliability correlation and regression. As stated in Chapter Three, for there to be a high correlation between scorers, or a similarity in the scores given by different raters for the same student, the r score should fall between .5 and 1; a medium correlation falls between .3 and .5; whereas a low correlation is below a .3 with a 0 signifying no correlation (Bluman, 2013, pp. 531-539). It is important to note most students were given multiple raters in the fall of 2014; however, due to lack of interest in scoring more assessments, the Assistant Dean of Accreditation and Faculty Development assigned assessments to fewer raters so there was a lack of multiple scorers for one student. Therefore, data for inter-rater reliability was only compiled for the fall. Figures 3-6 displayed the results for Tasks 1-4 the following hypothesis:

Hypothesis 3: There will be a difference in MoPTA scores given for the same teacher candidate by university supervisors and faculty.

Null Hypothesis 3: There will be no difference in MoPTA scores given for the same teacher candidate by university supervisors and faculty.

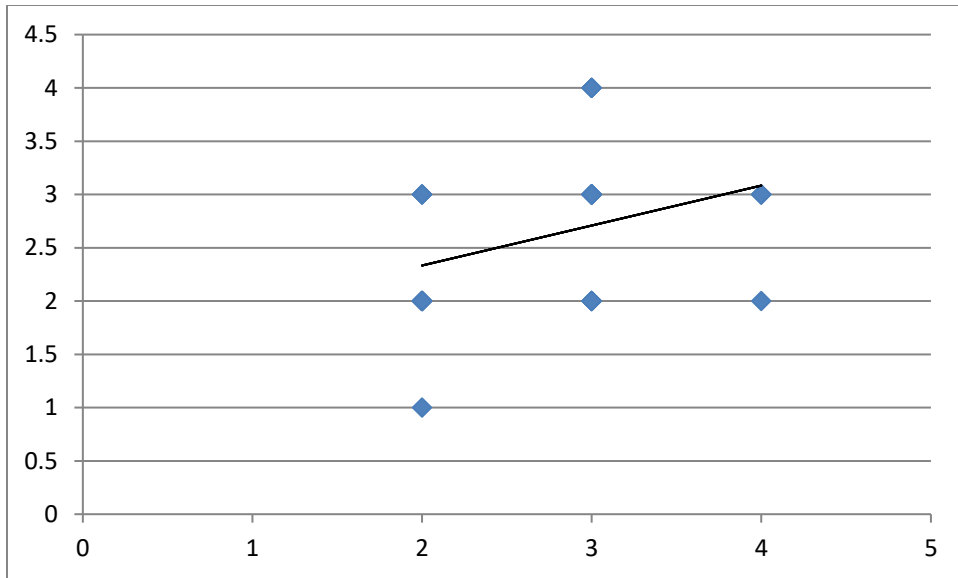


Figure 3. Inter-rater reliability for Task 1. Pearson ρ correlation coefficient for inter-rater reliability. $N=103$; $r=0.347$; $p=0.0003$

The scatter plot revealed the averages of scores between raters. If the scores were similar, the plots would gather around the regression line; however, since the plots were in various points on the graph and $r=0.347$ this signified a medium to low correlation in inter-rater reliability. Therefore, the null hypothesis was rejected and there was a difference between scores of one student as given by at least two different raters. Note in Table 4 it clarified 132 students completed Task 1, although inter-rater reliability examined a population of 103 in Figure 3. The difference of 29 was a result of only one rater for that student.

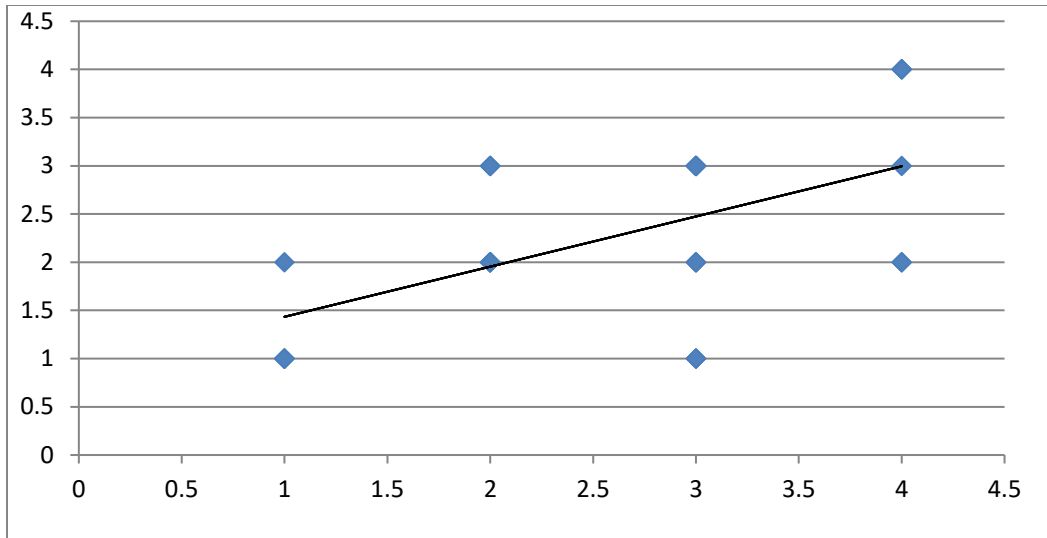


Figure 4. Inter-rater reliability for Task 2. Pearson ρ correlation coefficient for inter-rater reliability. $N=40$; $r=0.553$; $p=0.0002$.

The scatter plot revealed the averages of scores between raters. If the scores were similar, the plots would gather around the regression line; here, many of the plots were in similar points on the graph and $r=0.553$ signified a medium to high correlation in inter-rater reliability. Therefore, the null hypothesis was not rejected and there was not a difference between scores of one student as given by at least two different raters. Note in Table 4 it clarified 51 students completed Task 2, although inter-rater reliability examined a population of 40 in Figure 4. The difference of 11 was a result of only one rater for that student.

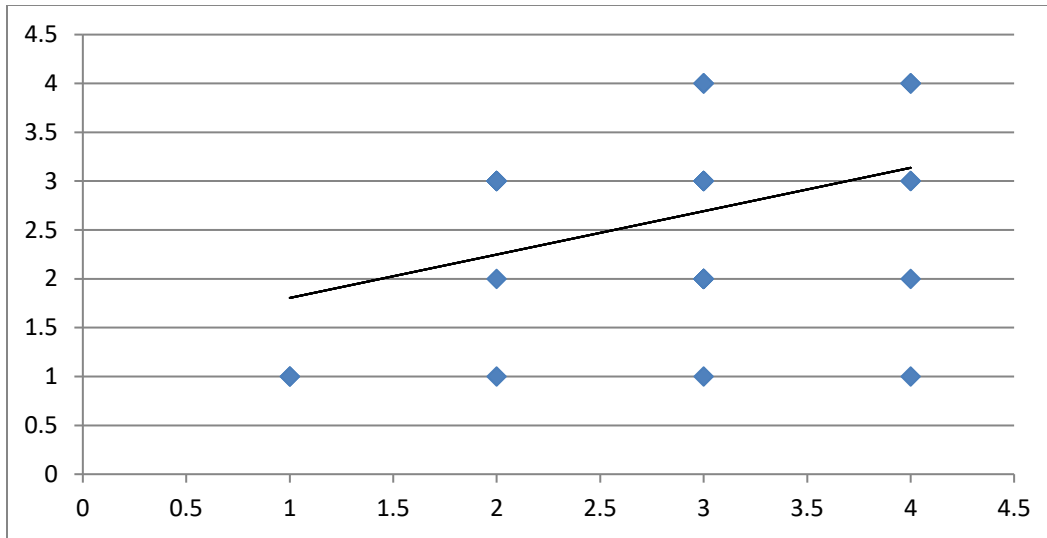


Figure 5. Inter-rater reliability for Task 3. Pearson ρ correlation coefficient for inter-rater reliability. $N=32$; $r=0.382$; $p=0.0310$.

The scatter plot revealed the averages of scores between raters. If the scores were similar, the plots would gather around the regression line; however, since the plots were in various points on the graph and $r=0.382$ this signified a medium to low correlation in inter-rater reliability. Therefore, the null hypothesis was rejected and there was a difference between scores of one student as given by at least two different raters. Note in Table 4 it clarified 53 students completed Task 3, although inter-rater reliability examined a population of 32 in Figure 5. The difference of 21 was a result of only one rater for that student.

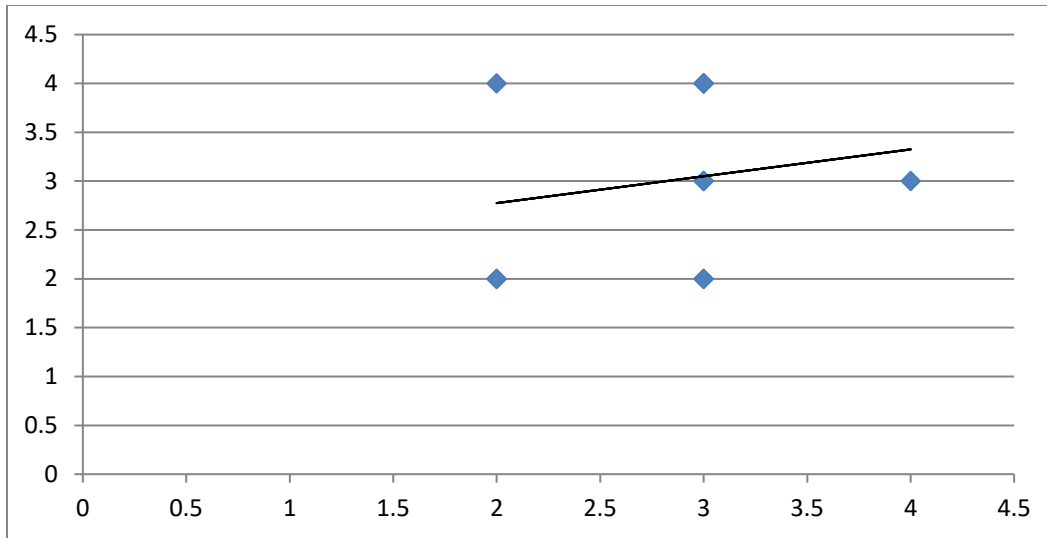


Figure 6. Inter-rater reliability for Task 4. Pearson ρ correlation coefficient for inter-rater reliability. $N=11$; $r=0.185$; $p=0.5860$.

The scatter plot revealed the averages of scores between raters. If the scores were similar, the plots would gather around the regression line; however, since the plots were in various points on the graph and $r=0.185$ this signified a low correlation in inter-rater reliability. Therefore, the null hypothesis was rejected and there was a difference between scores of one student as given by at least two different raters. Note in Table 4 it clarified 30 students completed Task 4, although inter-rater reliability examined a population of 11 in Figure 6. The difference of 19 was a result of only one rater for that student.

Null Hypothesis 4

This data was analyzed to investigate if there was a difference between the scores of elementary education students and secondary/K-12 education students, for each task.

Hypothesis 4: There will be a difference in MoPTA scores between elementary teacher candidates and secondary/K-12 teacher candidates.

Null Hypothesis 4: There will be no difference in MoPTA scores between elementary teacher candidates and secondary/K-12 teacher candidates.

The purpose of this test was to determine if students in the elementary education program performed at a higher level than those in the secondary/K-12 program. The outcomes could reveal if the curriculum of the elementary or secondary/K-12 program produced a higher score or if the program made no impact upon the score received on a task. Since the Fall 2014 MoPTA did not require students to complete which program they were enrolled, data was not compiled; however, due to this oversight, the Assistant Dean of Accreditation and Faculty Development changed the data input for the spring so data compiled in Foliotek was separated by the Coordinator of Elementary Teacher Candidates and the Coordinator of Secondary/K-12 Teacher Candidates, therefore the data was accessible. This change was made to accommodate the coordinators so they could view their individual students by their field, so it was easier for the elementary coordinator to filter data to find her specific students. Individual scores for elementary education students and Secondary/K-12 students for each task were entered into the ANOVA calculator, and the following tables displayed the results. Table 12 shows the overall results from the ANOVA test, which displays the number of students who took each task, by elementary or secondary/K-12 program, as well as the average of the scores.

Table 12.

Tasks 1-4: Elementary (EL) v. Secondary (9-12)/K-12 Students

Tasks	Count	Sum	Mean	Variance
Task 1: Elementary	77	210	2.73	0.7536
Task 1: Secondary/K-12	79	208	2.63	0.7481
Task 2: Elementary	36	88	2.44	0.8825
Task 2: Secondary/K-12	24	57	2.38	0.7663
Task 3: Elementary	29	69	2.38	0.6010
Task 3: Secondary/K-12	29	75	2.59	1.1084
Task 4: Elementary	12	37	3.08	0.8106
Task 4: Secondary/K-12	21	65	3.10	0.5905

Note. Task 1: $n=61$ (EL), $n=67$ (9-12/K-12); Task 2: $n=26$ (EL), $n=19$ (9-12/K-12); Task 3: $n=23$ (EL), $n=22$ (9-12/K-12); Task 4: $n=12$ (EL), $n=21$ (9-12/K-12). In some instances one student's task was scored by multiple raters, thus the "count" in the table is higher than the population (n).

Table 13 closely examined the difference of means between the elementary or secondary/K-12 program within each task using the Scheffe test, a test used with ANOVA when different sample sizes were used.

Table 13.

Scheffe Test: Elementary (EL) v. Secondary (9-12)/K-12 Students

Tasks	F_s	F_{crit}	Significant Difference?
Task 1: EL vs. 9-12/K-12	0.4624	3.903	No
Task 2: EL vs. 9-12/K-12	0.0830	4.007	No
Task 3: EL vs. 9-12/K-12	0.7262	4.013	No
Task 4: EL vs. 9-12/K-12	0.0016	4.160	No

There were no differences between the means of the tasks; therefore, the null hypothesis was not rejected.

Research Question 1

How do university supervisors perceive the process of evaluating teacher candidates' MoPTA tasks through the online portfolio system Foliotek?

Overall, university supervisors and faculty appreciated the experience of viewing and scoring students' work through the online portfolio system Foliotek. One stated she

“Liked having the insight in validating what I’m teaching” while another stated it was “Very telling to me to find [out] where student is now at the end of the program and progress showing.” Many also thought the experience was self-reflective in “Seeing pieces that were not addressed to evaluate own teaching of those areas” and the relation between the content of the tasks and the content of the class the university supervisor or faculty member taught.

University supervisors and faculty also enjoyed the scoring sessions in general. Many commented on the benefits of having multiple raters and discussing the tasks as well as examples. One stated “The discussion about each part is so valuable” as another agreed “The overall discussion about where we are and where we need to be is very valuable.” Another participant continued, “The opportunity to learn the comments of my colleagues was great. Also, it was good to receive feedback or suggested scores.” Another was thankful for the opportunity the scoring sessions provided: “Great discussion from the group. Really appreciate group input on the tasks. Valuable information and thoughts were expressed today. Thanks to both of you (Assistant Dean of Accreditation and Faculty Development and the researcher) for your energy and efforts.”

Many commented how nice it was to have a variety of members of the university there, from adjuncts to professors. In the past, supervisors were the only ones who scored portfolios during student teaching and it was done in isolation, so the scoring sessions appeared to be a welcome change. There are not many opportunities for all of the supervisors of student teachers to come together in a group, and all were grateful for the time and many of the positive attitudes they maintained throughout the sessions. As one participant noted, “It was nice to gather with peers!”

Furthermore, many learned a great deal regarding students' level of preparedness and how to score MoPTA properly. One commented that the "Scores varied a great deal 1-4," which coincided with many participants grateful for the multiple raters and discussions over the examples. A participant stated, "I valued the duplicity in scoring. Measuring my judgments against other educational opinions helped me feel more secure in my evaluations" and another commented "It was nice to have exemplars and to have (the Assistant Dean of Accreditation and Development and Coordinator of Secondary Teacher Candidates) there to answer questions." A similar comment agreed in regard to extra help on hand:

It was helpful to have us all together, studying the samples, and getting a feeling for a "2" or a "3". It was necessary to have (the Coordinator of Secondary Teacher Candidates and the researcher) here because just navigating the computer sites was often a challenge, and (the Assistant Dean of Accreditation and Faculty Development) can't help everyone at once. I think (the Assistant Dean of Accreditation and Faculty Development) did an excellent job!

There was a useful suggestion regarding examples in that "A true '4' [student examples receiving the highest possible score] I would like to see more." Others clarified "It helps me to understand better how to score these and what to look for in reading through them" a similar remark was "The more I do the more confident I feel because I am more familiar with the prompts. I am more confident that I'm not forgetting an important part of a prompt when reading the response." Another scorer agreed, "The more familiar I am with the tasks, the easier it is to score" and in agreement a fellow scorer felt "Very prepared. I like that we meet and discuss one example. It really helps me put things in

perspective.” A participant also concurred “I am much more comfortable scoring now after a couple of sessions-I may also be more critical.” Another reflected it was “Easy to score once you get the pattern down” and “Scoring MoPTA makes one more aware/familiar w content and expectations.” Although one scorer said the “first session was rather challenging focusing on task requirement and matching with student responses” she later clarified “the training sessions have been very beneficial providing the comfort and knowledge base to score task appropriately (based on the rubric).”

University supervisors and faculty also enjoyed viewing the tasks in the same manner as students completed the tasks. For example, some of the comments were it “Gave insight to what students know”; “What they are expected to do”; “What the tasks looked like,” and “It makes sense-answer the question; took out scare factor.” This next comment was in agreement by stating she was “More confident talking to students about it.” A participant reflected positively,

I was able to view student writing, and experience the MoPTA just as the students do. It was good to see that teacher candidates are gaining experience in preparing for their classrooms by gaining understanding about the socioeconomic status and other details about their school districts.

Many others concurred with this positivity with “It was very enlightening to see how much impact the pre-service classes have had on students as they answered the questions and gave examples” and “Gaining Insight into what our students need to know and be able to do; provided a frame of reference when discussing MoPTA with students.” In agreement,

this helped to know what my students were asked to write about in this new required format. As I am observing them I can now focus on the content of the tasks to help them be aware of the lessons and pupils they might be able to use in their upcoming writings.

Furthermore, as the scoring sessions continued (i.e., Tasks 2-3, Task 4 sessions), participants noticed improvement in the students' writing "quality seemed better this time. All parts of the questions (prompts) were addressed on those I read."

The participants also felt more prepared after the scoring sessions, as many commented "I feel more prepared and less unsure of myself" and one felt "Much more prepared." Another comment on preparedness reflected it was

Very necessary to have this kind of setting. Allows us to get an idea of what is the parameters for a 2-3 (hopefully not a 1). Much more prepared than last time Scoring the students' tasks on-site provides a good "refresher" for all of us each semester. Like any learning experience, I feel more prepared with each new session. I'm not an expert, by any means, but (the Assistant Dean of Accreditation and Faculty Development) does a good job of trying to keep us all "on track."

Another showed a similar sentiment: "The scoring sessions have helped to better understand the process and the scoring guide. It has been time well spent. I was more comfortable in the scoring during the second session. It helped me to review and collaborate"; and "it is important that we score these together for the support."

However, there was some negativity associated with scoring MoPTA. One participant stated, "Apparently we have no choice in the matter. It has been decided by

DESE as the way to go. We have to go with the flow and change is always hard whether young or old.”

Another drawback was the problems with logging on the computers and/or Foliotek and navigating the system. Due to these technical difficulties, many university supervisors and faculty reflected on the problems with using the computer and getting logged on to Foliotek as opposed to using it as scoring system. For example, when computers would lock up, one tried to use Foliotek password to get on the computer. Also it took about 45 minutes to get everyone logged in for Task 1 in the fall; however, in the spring it only took about 15-20 minutes for the Task 1 scoring session. Even though it was better than the fall, scorers still had trouble remembering usernames and passwords. Approximately 10 out of 26 of the scorers had trouble with computers. One faculty member commented, “If this doesn’t work, I’m going home.” For the Task 4 video component, there were many technical difficulties such as no sound, the video was imbedded upside down, or the video simply would not play. The Assistant Dean of Accreditation and Faculty Development suggested the raters be kind with scoring due to these problems. In other words, telling the scorers to keep in mind there may have been problems with the video component so to not have that impact the score of the overall writing of the student. One scorer summed up both of these issues,

The online Foliotek system actually works pretty well. It’s easy enough to navigate and there’s plenty of opportunity to write comments in the spaces provided. I do notice that it “drops me” occasionally when I have taken a little extra time to read a student’s response. If this happens toward the end of an evaluation, all ratings and comments are lost. Possibly I’m not privy to ways to

avoid this. One concern I have about this system is the Task 4 requirement of videotaping a student while teaching. I've only evaluated one Task 4 video, but I found it to be very "one dimensional," at best. Sure, I can hear and observe the student teacher providing a lesson, but I don't get to hear much feedback from students. With teaching and learning being a two-way street, I need to see much more student interaction to make the video worthwhile. If the purpose of the video is for prospective employers to view the candidate in action, I doubt they will get much out of it.

Despite the technical difficulties, overall the participants valued the scoring sessions and felt Foliotek was a useful tool for evaluation. As noted, "Accessing Foliotek is sometimes challenging but after logging in, it is user friendly."

University supervisors and faculty also commented on the process. Raters changed their views and debated scores during the scoring sessions. Overall university supervisors and faculty wanted more specific detail from students and a more clear explanation of baseline data in regards to pre-test and post-test (before and after) and its application in the classroom. They also suggested students labeling their responses a, b, c, d in response to each lettered questions within the textbox would lead to more clear guidance to the location of the answer. There was also concern about the artifacts overall since some student teachers may not have control over the artifact used, especially with Task 2; in terms of baseline data, that may be pre-determined by the district or school.

Research Question 2

How did faculty change the content of their lessons after evaluating teacher candidates' completed task(s)?

University supervisors and faculty were asked to reflect on changes they made in their classroom after examining and scoring the tasks. This question only applied to those teaching classes, as opposed to those whose only job for the university was supervising student teachers. Many had not made any, especially in the fall, but others did make adjustments. One stated she was “Teaching community as a triangle” and “showing visual of what the community looks like.” Others stated they were “Already adding new activities to reflect tasks” and “Already adding standards.” One suggested “Sharing scoring results with our students.” Many scorers noted the need for more writing instruction and assistance as they commented “Looking at student writing ability”; “Conferencing”; “Recognizing vocabulary and learning correct definition.” Furthermore, one started using the MoPTA template and artifacts in the classroom “We use the MoPTA lesson format, our students are given assignments that are similar to MoPTA tasks, and we are assigning more rigorous writing tasks.” Another stated “Interest inventories have always been recommend during instruction to know students and now a required component of Task 1.” Similarly “I am putting much more emphasis on identifying teaching strategies and expressing a rationale for choosing one over another” and in agreement, “I am using the MoPTA reflection template as a basis for writing reflections on lab/peer/microteaching.”

More included using MoPTA documents and vocabulary, “I am also using the Summative Assessment document to give feedback about their teaching” and “Using academic language and intentionally modeling, teaching and using that language with our students.” Another used the context of MoPTA by “Embedding more opportunities to practice similar tasks” and “We are already requiring more writing” along with “I try to

reinforce the use of appropriate terminology (e.g. instructional strategies, classroom management strategies).” Furthermore, “If everyone was paying attention people should have tweaked their class. I know I have by adding vocabulary and writing examples to my classes.” During one discussion participants asked if they could practice MoPTA in their classes, which all agreed was a good suggestion if there was ample time. Another concern was university supervisors and faculty expressed that current courses do not address content covered in textboxes. Some students included more details than necessary or the opposite, writing was not specific enough. However, overall university supervisors and faculty stated they were proud of their students and excited to make changes in their classroom to better prepare them for the classroom. Furthermore, all of the suggestions noted above are in direct alignment with MoPTA requirements and will greatly benefit teacher candidates.

Research Question 3 and Research Question 4

How do faculty and university supervisors perceive the teacher candidate preparation process (at this particular university)?

After participating in the piloted program with accompanying training, how do faculty and university supervisors perceive teacher candidates’ level of preparation for full MoPTA implementation in the fall of 2015?

Since many university supervisors and faculty focused on needed improvements for student teachers instead of reflecting on the teacher candidate preparation process, this research question was combined with Research Question 4. Therefore this addressed teacher candidate preparation overall. Some noted for Task 1 that there was “Not much interacting with families of students in letter to parents,” which was one of the required

artifacts. Also noted was “Students did not address building a community of learners” and “Writing skills, include more writing in all classes.” There were many comments on the lack of writing skills and the need for more academic language in the teacher education courses. A participant noted, “The students need additional support on how to respond to prompts to specific details and supporting evidence.” One further suggested,

More writing in coursework in the focal point. However, more experiences as listed above and determined by Department of Teacher Education PLC

Have students in classroom and teaching lessons for at least two semesters before student teaching. 1 day per week (or 2 half days per week) would be good.

What is being put in place now will be very helpful to our student teacher candidates. Every once in a while I get a student who cannot write and this still very concerning to me.

Another agreed with the importance of context to writing, “The students need more practice in applying the standards to classroom situations.” Further comments regarding writing suggested: “professional writing”; “rationales could be stronger. Students need to reflect on the ‘why’”; “making sure students elaborate with examples and good rationales. They should also keep their audience in mind when writing.”

Others claimed the students lack of knowledge to answer the tasks appropriately such as “How to make modifications for at risk populations”; “Resources-what they are, what tools, etc.”; “Direct writing to prompts, using information from a website (such as the demographics from DESE), and differentiation. Teacher candidate understanding of the impact(s) of demographics and other factors in student achievement.”

Further content that needed to be addressed: “The areas of deficit I see are writing skills and classroom management”; “They do not know what a ‘teaching strategy’ is and what a ‘learning activity’ is and how to express why one strategy would be chosen over another”; “making modifications for students”; “Classroom Management and Differentiation”; in Task 2 I noticed some of the candidates spoke about how they would change their teaching strategies (which in some cases was good formative assessment) instead of how they would modify the selected assessment plan”; “Classroom teaching and management needs to clearly define behavior management objectives.”

Comments also referenced the communication between teachers and parents, as well as the use of technology in the classroom. “Working with families and identifying the components of two-communications”; “Underscoring that not all families have access to the Internet and an open invitation to email the teacher is not really an invitation when a parent doesn’t have access to email.” In agreement, “It is evident students are not clear on critical thinking, inquiry based learning, and what constitutes a strategy.” Besides the core skills and ideas within MoPTA, one commented on students understanding and using different types of technology in the classroom. “More technology besides Smart board, such as a document camera and data gathering/storing programs and how results dictate what we do next.” Many stated in discussion that students would often write they did not use technology when the task textbox specifically asked for a technology to be used as a resource.

Some participants thought test-taking type skills were needed, such as “Answering questions asked-make sure answer correlates completely with question” and “Make sure they read directions closely. Make sure they answer question that was asked

before submitting answer.” The textboxes of some of the tasks proved quite lengthy with various prompts needing responses. Many mentioned students needing more exposure to “cultural diversity” in order to adequately answer some of the textboxes as well as familiarity with research-based practices and “Child development pathways as related to great instruction/classroom management methodology.”

University supervisors and faculty also offered advice for making needed changes. One participant stated,

That needs to be decided by pulling people together and see what changes have been made by different classes or [the] easiest way [is to] have each teacher send a quick report by list of what actual changes have been made. Thus [university administration] will be able to see if enough changes have been made.

Another agreed, “Throughout this training, discussion about numerous topics has helped professors know what to emphasize in classes. We need to meet and continue to dialogue. Maybe each semester next year and then end of the year after two years.” One suggestion, which may help this was, “Develop a PLC for faculty.” Although PLCs for faculty do exist in the School of Education, this was referring to a specific PLC working directly with MoPTA. In continuation, some comments reflected on specific courses and activities. “The methods classes need to be certain to include methodology that pertains to pre-kindergarten classes since they are becoming universal in school systems”; and “Secondary methods needs to give experience in these in the classroom,” in concurrence “Reading about it, hearing it articulate and discussing it are not enough if there is no practical experience.” Experience was also stressed in this commentary, “Students would benefit from more pre-student teaching classroom experiences. If students taught more

lessons and spent more time in classrooms before student teaching they could draw from these experiences when answering prompts.” Along with experience, was feedback on the seminar course, which took place during the student teaching semester:

Seminar has been great for the students. Might need more emphasis to the student to review videos, take the seminar class and include a practice MOPTA in their coursework, somewhere along the line. Possibly a step by step review of each task in a class prior to student teaching.

Another suggestion for courses clarified “What seems very evident is that we need to make a vocabulary list and embed these words in all methods classes”; “All classes should contribute to successful completion of the MoPTA tasks”; and a course suggestion was, “An additional written communications class might be helpful in the teacher prep program. If nothing else, it might re-employ all those English majors who are still looking for a job.” A further proposal, “(one faculty member) suggested we keep a list of effective classroom management (and academic strategies) we see our student teachers using. We could give those ‘lists’ to new student teachers as a ‘reminder’ of options available.” A comment relating to all the course proposals was one technique to be used across the curriculum by using “open dialogue in classes where students must defend, articulate, and support their techniques, rationale, and strategies.”

Overall, participants felt there was more to be done to prepare teachers but at the same time were hopeful for the future. “In my experience, the students are more prepared the last year than previously. I feel that the rigorous preparation in their education courses has greatly influenced this.” In agreement:

In general, the students seem to be well-prepared. There are a few slackers occasionally, but for the most part, these young people have a good grasp of what it takes to be a “beginning” teacher. They could use some help with grammar though. Having graded numerous rationales and reflections, this appears to be a weakness in their academic preparation. Possibly it’s due to this generation’s love for texting and sending shortened messages, but when sending a letter home to parent, the parents must have faith that the teacher has good command of the English language. When other people read what we have written, they’re measuring our intelligence (from afar) based upon what they read.

Another concurred with praise, “I think we have a good plan in place and our students are being made aware of what they will need to know and what it takes to be an excellent teacher. It goes without saying that all programs and plans can always use tweaking.”

Research Question 5

After participating in the piloted program with accompanying training, how do faculty and university supervisors perceive their preparation for full MoPTA implementation in the fall of 2015?

Not many responded to this question, which the researcher presumed was due to it listed as the last question and on the back of the survey; so many participants may not have turned the paper over to see there was a fifth question. Although many felt comfortable by the end of the scoring sessions, there was also a slight fear of change. For example, one frustrated scorer commented during a discussion, “How did teachers let this happen? Let DESE decide that one test determines certification?” Another hindrance to preparing students was the weight of the cost on students’ shoulders. During one scoring

session, a universal gasp went across the room at the thought of paying \$90 if a student fails a task.

Besides this, since there were already slight changes to MoPTA from the fall to the spring, participants voiced concerns over more changes coming in the fall. Also, as teacher candidates have begun taking the new state assessments there is concern on the pass rate as stated “Weaknesses have been noted in science, math, social studies, and arts content areas of certification assessments (CBASE/MoGEA/PRAXIS II). What can be done differently to help prepare to pass assessments?” However, as seen in the previous research question summaries, they do feel the scoring sessions prepared them for the content of MoPTA and are ready to make needed changes in order to properly prepare teacher candidates.

Summary

This mixed-methods study showed improvements were needed in the areas of Task 2, Assessment and Data Collection to Measure and Inform Student Learning and Task 3, Designing Instruction for Student Learning. The feedback from scorers provided many suggestions for making this a reality. Inter-rater reliability also showed a low correlation, which suggested more consistency was needed in order for students’ tasks to be evaluated fairly and efficiently. Furthermore, program differences were noted between Task 1 and 3 in terms of graduate students scoring higher than undergraduate students. However, no differences were noted between elementary and secondary/K-12 teacher candidates. The qualitative data showed overall university supervisor and faculty confidence in preparing students for the MoPTA, although they realize there was much work to be done. The next chapter provides suggestions for university supervisors and

faculty, as well as teacher candidates in order to adequately implement and complete MoPTA in the fall of 2015.

Chapter Five: Discussion and Reflection

Overview

In order to evaluate the teacher education program for the state of Missouri, the researcher investigated the piloted MoPTA at a private Midwestern university. Through evaluating the piloted MoPTA program, this study aimed to address possible changes needed within the teacher education program itself in order to better prepare future educators not only for the assessment, but more importantly, the classroom. In order to evaluate the program, the researcher observed scoring sessions for Tasks 1-4 and analyzed feedback from the evaluators of the tasks (university supervisors and faculty) in the fall of 2014 and the spring of 2015. Furthermore, this study examined the scores received from each task during the piloted school year (fall of 2014 and spring of 2015). The researcher analyzed the scores for the following comparisons: Tasks 1-4 (i.e. Task 1 overall scores to Task 2 overall scores); undergraduate students with graduate students' scores; inter-rater reliability (comparing the scores of multiple raters for one student); and K-12, elementary, middle, and secondary teacher candidates. By completing quantitative analyses of the comparisons, the researcher hoped to accomplish the following: examine the task(s) where students excelled, and where students struggled; analyze specific objectives not addressed in the university classroom; investigate discrepancies and lack thereof between scores of multiple raters per one student; and provide feedback regarding the effectiveness of an undergraduate program versus a graduate program. Through this investigation of the piloted MoPTA, the researcher hoped to possibly pinpoint specific and necessary curriculum modifications at the researched university to adhere to state and national standards for educators.

Discussion

Null Hypothesis 1.

Through examining the results of the piloted MoPTA in the fall of 2014, Task 2 (Assessment and Data Collection to Measure and Inform Student Learning) was the most difficult task when compared to Task 4 (Implementing and Analyzing Instruction to Promote Student Learning). The results of this ANOVA hypothesis test could possibly assist teacher candidates in Task 2 preparation. Task 2 specifically addressed “Assessment and data collection to measure and inform student learning” (MODESE, 2013c, para. 1). Teacher candidates were evaluated on the evidence from tests given to students and how the data from the assessments were used to increase student learning (MODESE, 2013c). Due to the statistical analysis involved in data collection and the fairly new use of data collection by teachers to “measure and inform student learning” the researcher concluded that Task 2 scores were the lowest. However, noting these results possible improvement could be made within the university classroom to better prepare teacher candidates to effectively use assessments and data in the classroom in order to improve student learning. Furthermore, Task 2 could be moved to the last task to be turned in since it is the most difficult and student teachers may have more data to discuss after more time has been spent in the classroom. It is also important to note the degree of difficulty involved in judging the effectiveness of a program so soon after its implementation since it takes time to see real results.

Through examining the results of the piloted MoPTA in the spring of 2014, it was clear that Task 3 (Designing Instruction for Student Learning) was the most difficult task when compared to Task 4 (Implementing and Analyzing Instruction to Promote Student

Learning). The results of this ANOVA hypothesis test could possibly assist teacher candidates in Task 3 preparation. Task 3 specifically addressed “ability to develop instruction, including the use of technology, to facilitate student learning” (MODESE, 2013d, para. 1). At the undergraduate level, the classroom management course was split into two courses, both with an emphasis on the topic but with two specific frames: one focusing on technology and the other on differentiation. However, not all student teachers were undergraduates, and even those that were not did not have all split classes. Similar to Task 2, teacher candidates were evaluated on the evidence from assessments given to students and how it affected instruction strategies and implementation to increase student learning (MODESE, 2013d). Noting these results, possible improvement could be made within the university classroom to better prepare teacher candidates to effectively use assessments to effectively design and implement instruction in order to improve student learning.

Performance on MoPTA.

There was a clear difference between scores for Task 2 and Task 4 in the fall of 2014. Task 2 (Assessment and Data Collection to Measure and Inform Student Learning) scores were significantly lower than Task 4 (Implementing and Analyzing Instruction to Promote Student Learning). This clearly revealed the need for specific instruction at the university level for the use of data in the classroom for instructors. As cited in Chapter Two, when data regarding teacher candidates was analyzed and shared, it led to a successful program with clear areas to improve and meet expectations (Margolis & Doring; Peck & McDonald, 2013). Which led the researcher to question, why should data sharing between university supervisors and teacher candidates vary from K-12

instructors sharing data with their students? Just as university supervisors are trained on accessing and analyzing data, so should K-12 teachers to know how to work with the numbers received from formative and summative assessments.

Student teachers also scored lower on Task 3 in comparison to Task 4 in the spring of 2015. Task 3 (Designing Instruction for Student Learning) proved to be difficult in opposition to Task 4 (Implementing and Analyzing Instruction to Promote Student Learning). As mentioned in the Literature Review, CCSS had students explain their reasoning behind problem solving math equations and TPA had students make connections between philosophies presented in their college courses and their own teaching strategies used in the classroom (CCSS Initiative, 2014; Margolis & Doring, 2013). This is the type of understanding Task 3 aimed to address. The researcher concluded, the development of critical thinking skills were equally important for both K-12 students and teacher candidates who will soon be full-time teachers in their own classrooms.

Null Hypothesis 2.

Through examining the results of the piloted MoPTA in the fall of 2014, the graduate students scored observably higher on Task 1 (Knowledge of Students and the Learning Environment) and Task 3 (Designing Instruction for Student Learning) than undergraduate students. The results of this ANOVA hypothesis test could possibly assist undergraduate teacher candidates in preparing for Task 1 and Task 3. Both of these tasks addressed the importance of knowing students' prior knowledge and experiences and adapting instruction to fit those needs (MODESE, 2013b, 2013d). Perhaps there was learning involved at the graduate level that could benefit undergraduates. Collaboration

between the two levels of students could lead to improved scores for undergraduate students. However, it may also be the experiences of the graduate level students that led them to higher scorers regarding Task 1 and Task 3. Either way, it is beneficial for the university to examine how the graduate program addressed prior knowledge and instruction design in juxtaposition to the undergraduate program in order to improve all teacher candidates' performance on the MoPTA.

Undergraduate v. Graduate Programs.

A significant difference was found between students in the undergraduate program versus the graduate program at this university. Graduate students scored higher on Task 1 (Knowledge of Students and the Learning Environment) and Task 3 (Designing Instruction for Student Learning); however, no difference was found between their scores on Task 2 (Assessment and Data Collection to Measure and Inform Student Learning) and Task 4 (Implementing and Analyzing Instruction for Student Learning). This showed a possible curriculum change needed at the undergraduate level to address applying knowledge of students to the lesson planning process and designing lesson plans. However, research also showed that collaboration was key to success in writing and student teaching experience (Robinson, 2014; Wichadee, 2013). The graduate program at this university offered courses in clusters, which could have led to more collaboration, hence higher scores on Tasks 1 and 3. Furthermore, the more experienced students may have felt more comfortable with implementing students' knowledge and designing instruction as opposed to less-experienced, undergraduate students. As for the lack of differences on Tasks 2 and 4, it may have been the sheer difficulty of application

with data for Task 2 and the convenience of using a video reflection for Task 4. More use of the critical thinking skill of application should be incorporated across the curriculum.

Null Hypothesis 3.

Through examining the results of the piloted MoPTA in the fall of 2014, it was clear there was a difference between scorers when evaluating one student's completed task. Inter-rater reliability was medium to low on Task 1 and Task 3, and low on Task 4. However, Task 2 yielded a medium to high inter-rater reliability. The results of this Pearson ρ correlation coefficient could possibly assist university supervisors, faculty, and trainers for the scoring of MoPTA in improving rater reliability. Furthermore, it could lead to an analysis of the positive inter-rater reliability for Task 2 as opposed to the medium to low reliability for the other tasks. An analysis such as this may improve training sessions and communication regarding protocol and collaboration regarding evaluating the tasks.

Inter-rater reliability.

Difference between scorers when rating one student's task was evident on Task 1, 3, and 4. Research cited previously discussed the difficulty of consistently rating student's work (Riggs et al., 2009); however, it was also necessary for moral and legal reasons (Pullin, 2013). As Pullin (2013) suggested, requiring students to pay to retake exams with a failing score, there needs to be accountability and validity for the university scoring the assessment. Since ETS will be scoring the MoPTA during full implementation, it should be suggested to them to stay within the legal guidelines outlined by Pullin. However, even though they will not be scoring the MoPTA in full in the future, further training for university supervisors and faculty should be developed to

score the tasks properly. This type of training could assist in other types of assessment scoring to keep validity high for students' scores. Furthermore, the reasoning of Task 2 scores revealed no difference and could have possibly been due to the overall low scores of that task where many students just received a 2 for completing the minimal, which may have kept many scorers consistent with incomplete and/or sufficient work.

Null Hypothesis 4.

Through examining the results of the piloted MoPTA in the fall of 2014, there were no observable differences between the scores of elementary education teacher candidates and Secondary/K-12 education teacher candidates. This clarified that a collaboration existed between the two programs at the university which should continue for the benefit of all teacher candidates.

Elementary teacher candidates v. secondary/K-12 teacher candidates.

No difference was found between student teachers seeking elementary certification and those seeking secondary or K-12 certification. A difference may have been expected due to the higher content level consistent with secondary training as opposed to elementary. For example, a high school English teacher may have had more training with writing due to the education and involvement with writing in the classroom, as opposed to an elementary education student. This result spoke to the consistency of curriculum across all levels of the teacher preparation program at this university.

Research showed (Cochran-Smith et al., 2011; Darling-Hammond, 2011; Sandholtz & Shea, 2012) that effective TEPs provide collaboration, consistent curriculum, and effective experiences that train educators for the classroom. It was clear this university implemented a TEP that was consistent across all levels of instruction, which could be

attributed to collaboration between faculty members from elementary and secondary education.

Participants' perceptions of scoring sessions and evaluating MoPTA online via Foliotek.

Overall, university supervisors and faculty found Foliotek to be an accessible and useful tool when evaluating MoPTA. Although technical difficulties occurred, once they learned the system they enjoyed reading and evaluating the tasks online. Many also found the process self-reflective of their own teaching; as one participant stated, "Seeing pieces that were not addressed to evaluate own teaching of those areas." Research showed that self-reflection, for teacher candidates and those supervising candidates, was imperative in the growth process (Gates, 2013; McVey, 2008; Tamer, 2014). Based on the participants' responses and the researcher's observations, it seemed that using this system allowed teacher candidates to self-reflect using the scorer's feedback, and the professor could self-reflect on what he or she was or was not doing in the university classroom.

Participants also appreciated the scoring sessions. Many commented the joy they felt by just being surrounded by other faculty members and adjuncts that they were not able to collaborate with often. This type of training could have easily been accomplished online, however the human contact not only proved beneficial but crucial to the success of the sessions. Research has proven the benefits of collaboration in all aspects of instruction (Darling-Hammond, 2011; Robinson, 2014). Therefore, types of training where instructors of student teachers discuss how they would score a task, what types of changes they wanted to implement in the classroom, concerns they had with MoPTA, and other issues and ideas, was a beneficial process. As one participant commented, "The

scoring sessions have helped to better understand the process and the scoring guide. It has been time well spent. I was more comfortable in the scoring during the second session. It helped me to review and collaborate.”

Changing content of lessons after scoring MoPTA.

During the beginning sessions instructors had not made changes due to the fact they had just been introduced to MoPTA; however, many started to make changes as they became more familiar with the assessment and the skills they found lacking in the student teachers they supervised. One stated “I am putting much more emphasis on identifying teaching strategies and expressing a rationale for choosing one over another.” Other participants agreed with clarifying terms such as teaching/instructional strategies, learning activities, differentiation, and other terms used throughout the tasks. Along with definitions, faculty also wanted to include more writing instruction in their courses or in the establishment of new courses. Research showed the importance of writing and critical thinking skills for future instructors and the need for universities to provide this type of preparation (Chung, 2008; CCSSO, 2011; edTPA, n.d.; Margolis & Doring, 2013; Robinson, 2014; Sandholtz & Shea, 2012). Based on participant feedback, many are ready to make the needed changes to fully prepare teacher candidates for their own classroom. As a participant stated, noted in Chapter Four, “Throughout this training, discussion about numerous topics has helped professors know what to emphasize in classes. We need to meet and continue to dialogue. Maybe each semester next year and then end of the year after two years.” In concurrence in discussing making modifications to required practicum experiences, “Reading about it, hearing it articulate and discussing it are not enough if there is no practical experience.” This as well as a participant’s

suggestion to develop a MoPTA PLC for adjustments in objectives in the teacher education classroom shows that many university faculty and supervisors are ready to implement necessary changes.

Participants' perception of teacher candidate preparation.

Similar to the above reflections on changes in courses, participants wanted to see more practice for students in regards to writing and terminology; however, they mainly wanted to see more context incorporated into their education. For example, one supervisor stated, "The students need more practice in applying the standards to classroom situations." Many other participants agreed stating the tasks had no meaning if the student teachers had no experience or were not able to apply what they had done in the classroom to the specific question the task was asking. For example, when Task 2 asked for application of baseline data and the teacher candidate had no experience with this type of data nor did his/her cooperating teacher, it was difficult to respond to that prompt. As cited earlier, Gewertz (2013) had similar concerns in her study of TPAs, as she expressed teacher candidates were not able to relate the experiences to the tasks. Many supervisors gave suggestions to incorporating more observations and teaching opportunities before the student teaching began, which was something also supported by researchers as effective (Cochran-Smith et al., 2011; Gewertz, 2013).

Participants' perception of full implementation of MoPTA.

Although there was some frustration with the assessment changes made by the state department of education (MODESE), many felt ready to implement MoPTA in the fall due to the scoring sessions and the training received. However, there was concern over the cost students would have to pay if they needed to take a task again; which added

stress to the supervisor's role in observing and assisting teacher candidates. Although registration for the tasks as well as final costs at the time of this study were not yet published, estimates included a total price of \$275 with additional monies being spent if a task or tasks were given a score of 0 or 1 or were not turned in by the deadline (MCATE, 2014; MEGA, 2014; MODESE, n.d.b.). Cavanagh (2013) also cited problems with the cost of tests and the reliance upon testing companies to certify teachers as opposed to the universities that trained the teacher candidates. Participants also expressed concern over more changes to come with MoPTA and other assessments to certify teachers. However, with continued support from administration at the university, they may have the tools needed to help teacher candidates whether any changes may occur.

Reflection on the Piloted Program

The piloted program began on September 5, 2014 and concluded with Task 4 in the spring of 2015. It involved approximately 60 university supervisors and faculty and a total of 276 teacher candidates that participated in the piloted MoPTA at this researched university. Overall, it was a success. There was positive participation and attendance at the scoring sessions, and teacher candidates submitted their Task 1 and other assigned task in a timely manner and were receptive to feedback. However, attendance at the sessions did dwindle at the end of each semester, with much lower participants in the spring as opposed to the fall; so much so that the Task 4 scoring session was cancelled in the spring. One of the drawbacks noticed during this pilot was the lack of technological skills among many of the faculty participants who scored MoPTA. Although memorizing a password for Foliotek or being able to navigate Google Chrome were not necessary to evaluate and supervise student teachers, it was an important aspect of scoring MoPTA. It

was clear many supervisors felt more comfortable scoring from home on a computer that memorized their passwords and where they knew how to navigate their Internet browser of choice; however, scoring was not the only element of the importance of familiarity with technology. The supervisors were perceived as the contact for teacher candidates when they need to upload their writing, artifacts, and, for some, a video to Foliotek. The researcher believed that supervisors should become more familiar and comfortable with technology overall. Besides this, every other element of the pilot was inspiring and thought-provoking. The researcher was amazed by the time many supervisors took to examine each and every textbox of a task and artifacts. Often times they would conference with each other over a score or ask the Assistant Dean of Accreditation and Faculty Development to look over their work before they submitted their score. The level of responsibility the supervisors felt towards their teacher candidates was heart-felt and earnest. It was enjoyable to work with the supervisors and the researcher felt the teacher candidates were set up for success with these supervisors assisting them through their student teaching practicum.

The students also impressed the researcher with their task writing. In agreement with the participants' discussion during scoring sessions and survey responses, overall teacher candidates had a grasp of the tasks, took them seriously, and showed excellent reflection and insight. In assisting some teacher candidates with uploading their tasks, many remarked how much they learned while writing and that the tasks really made them examine their teaching practices and their association with the learning goals. Although the tasks were time consuming, it was refreshing to experience teacher candidates who viewed the days and months of work as time well spent.

A final consideration is looking back on the researcher's inter-rater reliability analysis. Using only the mean scores from the tasks and the specific indicators of Tasks 3 and 4 and taking Cronbach's alpha into consideration, the overall results showed a weak inter-rater reliability. This did not show that the scores of the university supervisor and/or faculty member were wrong and the coordinator was correct, or vice versa. The results simply revealed a disconnect between scorers. The researcher believed this should be investigated and corrected to ensure reliability for the scores given to student teachers on their tasks to obtain certification. In this study to analyze the inter-rater reliability between university supervisors and faculty and their scores for the same student, the researcher used a Pearson ρ correlation coefficient for correlation and regression. The results of scores for the same teacher candidate with different scorers were inconsistent, but it was a highly difficult process to decide between a 2 and a 3 or even a 3 or a 4. Even with rubrics, collaboration, and scoring multiple tasks, it was difficult and even impossible at times to say with certainty this was the absolute correct score. Also, having anchor papers from ETS would have been helpful. As scorers requested "a true 4" it was hard to accommodate without examples from the scoring company. As research cited by Riggs et al. (2009) with inter-rater reliability suggestions to improve consistency, it was still hard to use logic as opposed to going with "your gut" so to speak. It also seemed to be a challenge for university supervisors to be objective; even though the tasks were not labeled by names but by codes, many of the supervisors could tell which student it was and would state "I know she knows this!" but yet did not provide a clear answer. These are problems that will never be solved simply because we are human. However, it is important to note just as there is an element to teaching that cannot be measured, this also

held true when scoring the MoPTA. This by no means is to say there should not be evaluation measures such as MoPTA or that they should not be scored by supervisors, but it is necessary to keep these aspects in mind while scoring and examining scores. Our humanity ensures imperfection; yet there are recommendations and implementations that should be considered to increase reliability.

Recommendations for Program

The researcher has recommendations for the testing company (ETS), the teacher education program for this university, and for other universities to implement. ETS should improve upon the rubrics they developed for each task. They were so lengthy (see Figure 2 in Chapter Three) and not easy to use during scoring. Also, anchor papers to show what a 4, 3, 2, and 1 scored papers look like would be beneficial for students, scorers, and professors alike. Furthermore, more examples of artifacts for student use would be beneficial especially for different subject areas. Those are difficult to be created and having examples ETS has scored in the past would serve as effective models. Finally, ETS should move Task 2 to the final task. It not only proved to be the most difficult, but many teacher candidates do not have accessibility to data until well into their experience. Some may not even have a cooperating teacher that uses data in the manner requested by Task 2. These items need to be taken into consideration by ETS.

Universities can also attempt to rectify the gaps seen in Task 2, Assessment and Data Collection to Measure and Inform Student Learning and Task 3, Designing Instruction for Student Learning. Both of these address the gap of the ability of students to use data from formative and summative assessments to address student needs and plan lessons accordingly. The university certainly has courses in place to address needs such

as lesson planning, teaching theorists to address various student needs, and the use of data in the classroom. However, the classes were not designed with the MoPTA tasks simply because they did not exist at the time the curriculum was written. Now university faculty can align the tasks, specifically the textboxes, and gaps in the current curriculum not addressed by the individual tasks. The researcher recommends individual assignments for groups of teachers to investigate the tasks in connection to the curriculum, and then write a proposal for curriculum changes based on these gaps. Furthermore, the concept of analysis is also addressed in each of these tasks and it is important for critical thinking skills such as analyzing data and learning theories to be placed at the forefront of educating not only teacher candidates, but also all students. The researcher believed this is a critical skill as a lifelong learner and students were being underserved by not incorporating analysis into their instructional strategies and learning activities.

Another needed change would be additional coursework on writing. This may be simply adding more writing to existing courses, but the researcher's recommendation is to isolate the course on writing so students may be able to use the course for the context of the writing they are completing. This will also make the course accessible to students across all levels of education and majors. As students continue to communicate more through online avenues and the written word, there has never been a more crucial time to stress the importance of choosing words with care and articulating sentences that are authentic and applicable to the situation at hand.

Furthermore, still dealing with curriculum and course changes, the researcher recommends a course to be offered which focuses on the tasks. The researcher actually created a course entitled "Writing for the MoPTA Tasks" during J-term (a course for the

weeks between fall and spring semesters). During the course the researcher went through each of the tasks and textboxes explicitly and students practiced writing one of the textboxes for each of the tasks as well as created artifacts to accompany the task. They also became more familiar with the terminology, such as knowing the difference between instructional strategies and learning activities, and differentiation and modifications. It was also practice for them to see and evaluate other students' writing and provide them feedback so they could improve their writing. The researcher perceived the course as successful and recommended implementation at the researched university, or elements of the course be included in existing courses.

Also, the researcher recommends the training sessions continue for supervisors of student teachers. Even though in the fall of 2015, when MoPTA is fully implemented, supervisors will only score Task 1 with MODESE scoring Tasks 2-3, it is still vital for supervisors to be able to continue to learn the aspects of MoPTA and share them with their teacher candidates. Staying involved is also a great way to relieve some of the stress they feel over the cost of the exam for teacher candidates. If they are continuing to learn and collaborating with their teacher candidate, it helps both of them feel at ease with the assessment and in the end produce a high score and even more important, a highly qualified teacher. Another benefit to the scoring sessions is continued work on inter-rater reliability. The only way to produce consistent results in scoring and to understand the difference between scores for one candidate is to collaborate and continue the conversation. The scoring sessions are an excellent platform for that type of professional development.

Finally, the researcher recommends as much support possible for teacher candidates. Similar to what was stated in the previous paragraphs, by offering curriculum changes, writing courses, and continued communication between the supervisor and candidate that is giving needed support. However, having a team of supportive individuals is equally important. The Assistant Dean of Accreditation and Faculty Development currently has a Blackboard shell with examples and much needed MoPTA information. This type of online service was viewed by the researcher as excellent and should continue, as well as setting up blogs and twitter feeds for teacher candidates to communicate with each other. Ultimately it is on the teacher candidate to complete the tasks and understand what is needed to become an effective teacher; however, this cannot be accomplished without the support of a knowledgeable and useful community of learners and instructors.

All of the above recommendations are impossible without thoughtful, committed, and well-trained supervisors of teacher candidates. It is imperative for this university, and others, to hire and train effective educators that are willing and able to embrace lifelong learning for themselves and the teacher candidates they will support. It was disheartening at times to observe some supervisors dismissing the use of technology as they assessed MoPTA on Foliotek. Many assumed it was fleeting and they would be assessing a different type of portfolio next year and thus did not see the benefit to learning a new system. Although this type of frustration is understandable in this ever changing world of technological advancements; however, this is no excuse for supervisors of student teachers to dismiss a current evaluation tool and refuse to commit to educating themselves on using the basic tools the Internet has to offer. Furthermore, these same

type of committed teacher candidate supervisors are also needed in the classroom. Many times we as educators do not practice what we preach. This researcher has sat in many a lecture hall throughout her years of college coursework, while the professor of education has lectured about the ineffectiveness of teachers lecturing students. In continuation of this point, the use of technology in the classroom and collaborative learning are two useful instructional methods that are barely used inside university walls; yet during education courses, the students are taught how effective they are, again in a lecture-type format. This paradox is inexcusable. Teachers who teach future educators must use the tools they are encouraging their students to use when they arrive in their classroom to student teach. Without truly modeling to future educators how to teach, we are not truly teaching teachers. Thus leading to an ineffective teacher workforce to instruct the K-12 students of the future.

Recommendations for Future Research

For the future, this study should continue through full implementation in the 2015-2016 school year. Performance levels on tasks, undergraduate versus graduate programs, inter-rater reliability, and elementary versus secondary/K-12 should all continue to be analyzed quantitatively; while further studying the results, implications, and recommendations through qualitative studies such as the survey instrument created by the researcher. Other universities should also complete a similar study to see where improvements should be made, and where they excel.

Other studies should include continuing to examine inter-rater reliability and various training sessions to improve consistency. As more and more assessments and evaluations move to the online format, this is easier to analyze and pinpoint where

scorers see different viewpoints when it comes to giving similar ratings. Collaboration is the key, so more opportunities to score together are imperative even though it can be done online. A future study involving scoring online, in a community environment similar to the scoring sessions, would provide feedback to the issue of scoring in isolation. It may be that scores could be more consistent if the scorers were able to collaborate throughout the process. Or, if the opposite is true, what can be done to identify the vital aspects of writing to be scored and how to score them as accurately as possible.

A further recommendation for future study is to gain feedback from teacher candidates regarding MoPTA and the other state assessments. Possible questions to ask could be how they felt about taking the MoPTA, if it prepared them for the classroom, and their opinion on the cost of the exam and the risk of not being certified based on the results of the exam. It could also be beneficial to look at performances of first year teachers and compare it to their MoPTA scores to see if it is predictive of teaching ability. No matter the method, as tuition costs keep rising, and tests increase in cost, it is vital to keep communicating with the students at all times and examining past and future data. They need to be listened to in order for them to have a productive university experience and become an effective instructor for tomorrow.

Conclusion

As the 21st century continues it is important to embrace the changing world of education by implementing teacher education programs that produce highly qualified educators for this new world of learners. Through incorporating critical thinking skills, evidence of work, written commentary, reflective practices, and effective resources in the

K-12 and university classrooms, students and instructors of those students will be well-trained, thoughtful, creative, and collaborative individuals. The world needs this type of student and instructor to create innovative ideas to fix the problems and inspire others.

The departments of education, statewide and nationally, can no longer rely on one professor at one university to evaluate whether a student is ready to be a full-time teacher in his or her own classroom. It takes collaboration for any change to take place and for implementation to be successful. This means using valid and reliable evidence from the classroom effectively. University supervisors can be the driving force behind this type of change by implementing programs that support assessments such as MoPTA to create consistency and a collaborative working environment for faculty and all instructors and supervisors of student teachers to make education programs challenging and productive.

References

- American Association of Colleges for Teacher Education. (2013). The changing teacher preparation profession. A report from AACTE's Professional Education Data System. Retrieved from https://secure.aacte.org/apps/rl/res_get.php?fid=145
- American Association of Colleges for Teacher Education. (2015). Participation map. *EdTPA*. Retrieved on June 12, 2015 from <http://edtpa.aacte.org/state-policy>
- Asplin, K. N., & Marks, M. J. (2013). Increasing the influence of university supervisors during student teaching. *Professional Educator*, 37(1). 1-11.
- Bill and Melinda Gates Foundation. (2010). Learning about teaching: Initial findings from the measures of effective teaching project. MET Project. *Bill & Melinda Gates Foundation*. Retrieved from http://www.metproject.org/downloads/Preliminary_Findings-Research_Paper.pdf
- Bluman, A. G. (2013). *Elementary statistics: A brief version* (6th edition). New York, NY: McGraw Hill.
- California Teacher Preparation Assessment. (2013). *Candidate handbook*. Retrieved from <http://www.ctc.ca.gov/educator-prep/TPA-files/CandidateHandbook.pdf>
- Carey, K. (2015). *The end of college*. New York, NY: Riverhead Books.
- Cassidy, C. (2015, April 18). Thousands of students opt out of Common Core tests in protest. Retrieved from PBS Newshour Website: <http://www.pbs.org/newshour/rundown/thousands-students-opt-common-core-tests-protest/>

- Caughlan, S., & Jiang, H. (2014). Observational and teacher quality: Critical analysis of observational instruments in preservice teacher performance assessment. *Journal of Teacher Education, 65*(5), 375-388.
- Cavanagh, S. (2013). Ed. testing industry sees rising demand. *Education Week, 33*(6), 1-17.
- Chung, R. (2008, Winter). Beyond assessment: Performance assessments in teacher education. *Teacher Education Quarterly, 35*(1), 7-28.
- Cochran-Smith, M., Cannady, M., McEachern, K. P., Piazza, P., Power, Ch., & Ryan, A. (2011). Teachers' education, teaching practice, and retention: A cross-genre review of recent research. *Journal of Education, 191*(2), 19-31.
- Common Core State Standards Initiative. (2014). *Read the standards*. Retrieved from <http://www.corestandards.org/read-the-standards/>
- Council of Chief State School Officers. (2011, April). *Interstate Teacher Assessment and Support Consortium (InTASC) Model Core Teaching Standards: A Resource for State Dialogue*. Washington, DC: Author.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Darling-Hammond, L. (2011, March 16). Teacher preparation: Build on what works. *Education Week, 30*(24), 36, 25-26.
- Dynarski, M. (2014, December 4). Moving teacher preparation into the future. *Brookings Institution*. Retrieved from <http://www.brookings.edu/research/papers/2014/12/04-chalkboard-dynarski-teacher->

prep?utm_campaign=Brookings+Brief&utm_source=hs_email&utm_medium=e
mail&utm_content=15171859&_hsenc=p2ANqtz-9M1Y4yAveOcCRa-
PBZW9W_uSZQanyvdKsShyzzLzM4GY1PtIooQW7yFD3FvLo3nVbQOphiE0i
HCZuo_q8eYa9YCLdQKA&_hsmi=15171859

EdTPA. (n.d.). About edTPA. *Overview*. Retrieved from <http://edtpa.aacte.org/about-edtpa#Overview-0>

EdTPA. (2014). Fees and payment options. Retrieved on June 13, 2015 from http://www.edtpa.com/PageView.aspx?f=GEN_OperationalFees.html

Educational Testing Service. (n.d.). Praxis performance assessments. The Praxis Series. Retrieved on June 13, 2015 from

Educational Testing Service. (2014). Frequently asked questions about the Praxis Series test. *The Praxis Series*. Retrieved from https://www.ets.org/praxis/faq_test_takers/

Fernandez, M. L., & Erbilgin, E. (2009). Examining the supervision of mathematics student teachers through analysis conference communications. *Educational Studies in Mathematics*, 72, 93-110.

Foliotek. (2014). Growing competent and passionate professionals. Retrieved on September 11, 2014 from Foliotek Website: <http://www.foliotek.com/>

Gates, B. (2013). Teachers need real feedback. *TED Talks Education*. Retrieved from https://www.ted.com/talks/bill_gates_teachers_need_real_feedback?language=en

Gewertz, C. (2013). Transition to online tests sparks fears (cover story). *Education Week*, 33(10), 1-15.

Goldhaber, D., & Cowan, J. (2014). Excavating the teacher pipeline: Teacher preparation programs and teacher attrition. *Journal of Teacher Education*, 65(5), 449-462.

Green, E. (2014). *Building a better teacher: How teaching works (and how to teach it to everyone)*. New York, NY: W. W. Norton & Company, Inc.

Hariston, G. (2013, August 21). Important changes in educator preparation. Memo from the Missouri Department of Elementary and Secondary Education: Office of Educator Quality. Retrieved from <http://dese.mo.gov/sites/default/files/GPAandMoGEAMemo.pdf>

Hariston, G. (2014, October 20). Missouri educator gateway assessments for 2014. Memo from the Missouri Department of Elementary and Secondary Education: Office of Educator Quality. Retrieved from <https://dese.mo.gov/sites/default/files/Educator%20Preparation%20Memo%2010-20-2014.pdf>

Jones, C. (2010). Editorial: Describing or debunking? The net generation and digital natives. *Journal of Computer Assisted Learning*, 26, 317-320.

Kithakarn, O. (2012). Using blogs to improve students' summary writing abilities. *Turkish Online Journal of Distance Education*, 13(4), 209-219.

Klein, J. (2014). *Lessons of hope*. New York, NY: Harper Collins.

Kubiatko, M. (2013). The comparison of different age groups on the attitudes toward and the use of ICT. *Educational Sciences: Theory and Practice*, 13(2), 1263-1272.

Lenhart, A., Arafeh, S., Smith, A., & MacGill, A. (2013, Spring). Writing, technology, and teens. *Pew Research Internet Project*. Retrieved from <http://www.pewinternet.org/2008/04/24/writing-technology-and-teens/>

Lindley, F. A. (2009). *The portable mentor: A resource guide for entry-year level principals and mentors*. Thousand Oaks, CA: Corwin Press.

Margolis, J., & Doring, A. (2013). National assessment for student teachers:

Documenting teaching readiness to the tipping point. *Action in Teacher Education, 35*, 272-285.

McVey, M. (2008). Writing in an online environment: Student views of "inked"

feedback. *International Journal of Teaching and Learning in Higher Education, 20*(1), 39-50.

Missouri Association of Colleges for Teacher Education. (2014). Concerns regarding the

MPTEA/MoPTA sequence of assessments. Retrieved from

http://associations.missouristate.edu/assets/macte/Concerns_Regarding_the_MPT EA__3_.pdf

Missouri Department of Elementary and Secondary Education. (n.d.a.). *Educator*

preparation. Retrieved from <http://dese.mo.gov/educator-quality/educator-preparation>

Missouri Department of Elementary and Secondary Education. (n.d.b.). *About*. Retrieved

from <http://dese.mo.gov/about>

Missouri Department of Elementary and Secondary Education. (2013a). Teacher

standards. *Missouri's Educator Evaluation System*. Retrieved from

<http://dese.mo.gov/sites/default/files/TeacherStandards.pdf>

Missouri Department of Elementary and Secondary Education. (2013b). *Task 1 rubric*.

Retrieved from

http://www.nwmissouri.edu/dept/peu/pdf/studentteach/mopta/Task1_Rubric.pdf

Missouri Department of Elementary and Secondary Education. (2013c). *Task 2 rubric*.

Retrieved from

http://www.nwmissouri.edu/dept/peu/pdf/studentteach/mopta/Task2_Rubric.pdf

Missouri Department of Elementary and Secondary Education. (2013d). *Task 3 rubric*.

Retrieved from

http://www.nwmissouri.edu/dept/peu/pdf/studentteach/mopta/Task3_Rubric.pdf

Missouri Department of Elementary and Secondary Education. (2013e). *Task 4 rubric*.

Retrieved from

http://www.nwmissouri.edu/dept/peu/pdf/studentteach/mopta/Task4_Rubric.pdf

Missouri Department of Elementary and Secondary Education. (2014, September, 23).

The transition: From MoSTEP and compendium to MoSPE and new certification requirements. Office of Educator Quality Webinar Retrieved from

<http://dese.mo.gov/sites/default/files/Webinar-Slides-9-23-14.pdf>

Missouri Department of Elementary and Secondary Education. (2015, February 23).

Missouri pre-service teacher assessment alternate: Task 4 rubric, draft. Retrieved on June 12, 2015 from

<https://dese.mo.gov/sites/default/files/Task%204%20B%20Rubric%20NTC%20%203-6-15.pdf>

Missouri Educator Gateway Assessments. (2014). Missouri performance assessments.

The Missouri Pre-Service Teacher Assessment (MoPTA). Retrieved on September 11, 2014 from http://mega.ets.org/pre_service_teacher_assessment/

Missouri Educator Gateway Assessments. (2015a). Missouri educator profile. Retrieved

on June 11, 2015 from

http://www.mo.nesinc.com/TestView.aspx?f=HTML_FRAG/MO006_TestPage.html

Missouri Educator Gateway Assessments. (2015b). Missouri general education assessment. Retrieved on June 11, 2015 from

http://www.mo.nesinc.com/PageView.aspx?f=GEN_Tests.html

Munby, H., Russell, T., & Martin, A. K. (2001). Teachers' knowledge and how it develops. In V. Richardson (Ed.), *Handbook of research on teaching*.

Washington, DC: American Educational Research Association.

Naidoo, A. (2012, June). Leading curriculum renew in a faculty of education: A story from within. *Perspectives in Education*, 30(2), 71-80.

Ongel, S., Capa, Y., & Vellom, R. P. (2002, April). *A study of the effectiveness of a preservice teacher education scheme*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Peck, C., & 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.

Private University. (n.d.a.). *Student teaching handbook*. Retrieved from

<http://www.xxx.edu/education/teachered/docs/handbook.pdf>

Private University. (n.d.b.). Teacher education department. *Undergraduate catalog 2015-*

2016. Retrieved from <http://xxx.smartcatalogiq.com/en/2015->

[2016/Undergraduate-Catalog/School-of-Education/Teacher-Education-](http://xxx.smartcatalogiq.com/en/2015-2016/Undergraduate-Catalog/School-of-Education/Teacher-Education-)

[Department](http://xxx.smartcatalogiq.com/en/2015-2016/Undergraduate-Catalog/School-of-Education/Teacher-Education-)

- Pullin, D. (2014). *Performance, value, and accountability: Public policy goals and legal implications of the use of performance assessments in the preparation and licensing of educators*. Washington, DC: Council of Chief State Officers and Stanford Center for Assessment, Learning, and Equity.
- Rantanen, P. (2013). The number of feedbacks needed for reliable evaluation. A multilevel analysis of the reliability, stability and generalisability of students' evaluation of teaching. *Assessment & Evaluation in Higher Education*, 38(2), 224-239. doi:10.1080/02602938.2011.625471
- Riggs, M. L., Verdi, M. P., & Arlin, P. K. (2009). A local evaluation of the reliability, validity, and procedural adequacy of the teacher performance assessment exam for teaching credential candidates. *Issues in Teacher Education*, 18(1), 13-38.
- Robinson, S. P. (2014, September/October). Managing and communicating change. *Principal* (pp. 25-27). Retrieved from <http://www.naesp.org/SeptOct14>
- Sandholtz, J. H., & Shea, L. M. (2012). Predicting performance: A comparison of university supervisors' predictions and teacher candidates' scores on a teaching performance assessment. *Journal of Teacher Education*, 63(1), 39-50.
- Sawchuck, S. (2014, October 21). Steep drops seen in teacher-prep enrollment numbers. *Education Week*. Retrieved from <http://www.edweek.org/ew/articles/2014/10/22/09enroll.h34.html?tkn=NLRFunrwoJuIV6Apnq0S4XV9VRXIC0LSjgY2&intc=es>
- The State of California. (2015). Teaching performance assessment. Retrieved on June 13, 2015 from <http://www.ctc.ca.gov/educator-prep/TPA.html>

- Tamer, M. (2014). The value of self reflection: With video cameras in hand, educators enhance their practice by watching themselves teach. *Usable Knowledge: Connecting Research to Practice*. Retrieved from <http://www.gse.harvard.edu/news/uk/14/10/value-self-reflection>
- Tellez, K. (1996). Authentic assessment. In J. Sikula (Ed.), *The handbook of research in teacher education* (2nd edition). New York, NY: Macmillan.
- Trucano, T. G., Swiler, L. P., Igusa, T., Oberkamp, W. L., & Pilch, M. (2006). Calibration, validation, and sensitivity analysis: What's what. *Reliability Engineering and System Safety*, 91, 1331-1357.
- United States Department of Education. (2003). Meeting the highly qualified teachers challenge. *The Secretary's Second Annual Report on Teacher Quality*. Retrieved from <https://www2.ed.gov/about/reports/annual/teachprep/2003title-ii-report.pdf>
- The Warner School of Education. ((2015, March 2). Study shows teacher candidates lack a clear understanding of edTPA. Retrieved from <https://www.warner.rochester.edu/newsevents/story/1528/>
- Westervelt, E. (2015, March 3). Where have all the teachers gone? Retrieved on June 12, 2015 from the National Public Radio Website: <http://www.npr.org/sections/ed/2015/03/03/389282733/where-have-all-the-teachers-gone>
- Wichadee, S. (2013, July). Improving students' summary writing ability through collaboration: A comparison between online wiki group and conventional face-to-face group. *The Turkish Online Journal of Educational Technology*, 12(3), 107-116.

Zeichner, K. M. (2003). The adequacies and inadequacies of three current strategies to recruit, prepare, and retain the best teachers for all students. *Teachers College Record*, 105(3), 490-519.

Appendix A: Survey Questions

1. Describe your experience scoring MoPTA.
2. After scoring students' responses, evaluate the teacher preparation program at this university.
3. Describe changes, if any, to be implemented to address the evaluation stated in question 2.
4. Evaluate the use of the online system Foliotek to score MoPTA.

Appendix B: Permission to use Survey and the Study Site

Leitsch, Patricia
 To: Elder, Robyne
 Attachments: [MoPTA-Remarks After Scoring-1.docx \(20 KB\)](#) [Open as Web Page](#)
 Monday, November 03, 2014 1:34 PM

- You replied on 11/3/2014 1:04 PM.

I have attached them for you.

Patricia

Elder, Robyne
 Sent from: [Send from my iPhone](#)
 Monday, November 03, 2014 12:57 PM

Thanks so much! I do not, but I think Dr. Beth mentioned she had them somewhere. I will see if she has them when she returns from her conference. I will type them up and send it your way.

Leitsch, Patricia
 Monday, November 03, 2014 12:40 PM

I have no problem with it. Do you have the first set of responses?

Elder, Robyne
 Sent from: [Send from my iPhone](#)
 Monday, November 03, 2014 11:44 AM

Good afternoon, Dr. Leitsch! I am writing my dissertation on MoPTA, and I wanted to see if I would be able to use the feedback/score's responses from your surveys as data for my paper.

Thanks so much,
 Robyne

sign out | Elder, Robyne

permission to use LU

Bice, Cynthia
 Wednesday, November 05, 2014 3:30 PM

All is approved as per Provost as long as IRB approves and we are not listed by name in the study.

Keep this email for future reference.

Good work!

Elder, Robyne
 Sent from: [Send from my iPhone](#)
 Wednesday, November 05, 2014 11:12 AM

Thank you!!

Bice, Cynthia
 Wednesday, November 05, 2014 11:03 AM

I am approving and sending on to provost...I will let you know ASAP.

Sent from my iPhone

Elder, Robyne
 To: Bice, Cynthia
 Attachments: [Elder MoPTA prospectus.docx \(30 KB\)](#) [Open as Web Page](#)
 Sent from: [Send from my iPhone](#)
 Wednesday, November 05, 2014 10:52 AM

Good morning, Dr. Bice. I wanted to request your permission to use LU as the study site of my dissertation. My topic is: A Mixed-Method Investigation of the MoPTA Pilot Program at a Private Midwestern University. I have attached my prospectus to provide more specific information regarding my study.

Thanks so much,
 Robyne

Appendix C: Observational Data from MoPTA Scoring Sessions

How observational data was gathered: while university supervisors and faculty were scoring MoPTA tasks using the online portfolio system Foliotek, the PI was walking around the room assisting with any questions or problems (as a part of her Graduate Assistant work); while doing so, she make observational notes on the following:

Scorers' behavior/feedback while operating Foliotek to access students' responses and submit students' scores (problems, positives, questions, etc.):

Negative behavior/feedback							Positive behavior/feedback		
1	2	3	4	5	6	7	8	9	10

Notes:

Scorers' behavior/feedback toward the students' responses:

Negative behavior/feedback							Positive behavior/feedback		
1	2	3	4	5	6	7	8	9	10

Notes:

Scorers' behavior/feedback toward MoPTA:

Negative behavior/feedback							Positive behavior/feedback		
1	2	3	4	5	6	7	8	9	10

Notes:

Scorers' behavior/feedback toward their own instructional strategies after reading students' response:

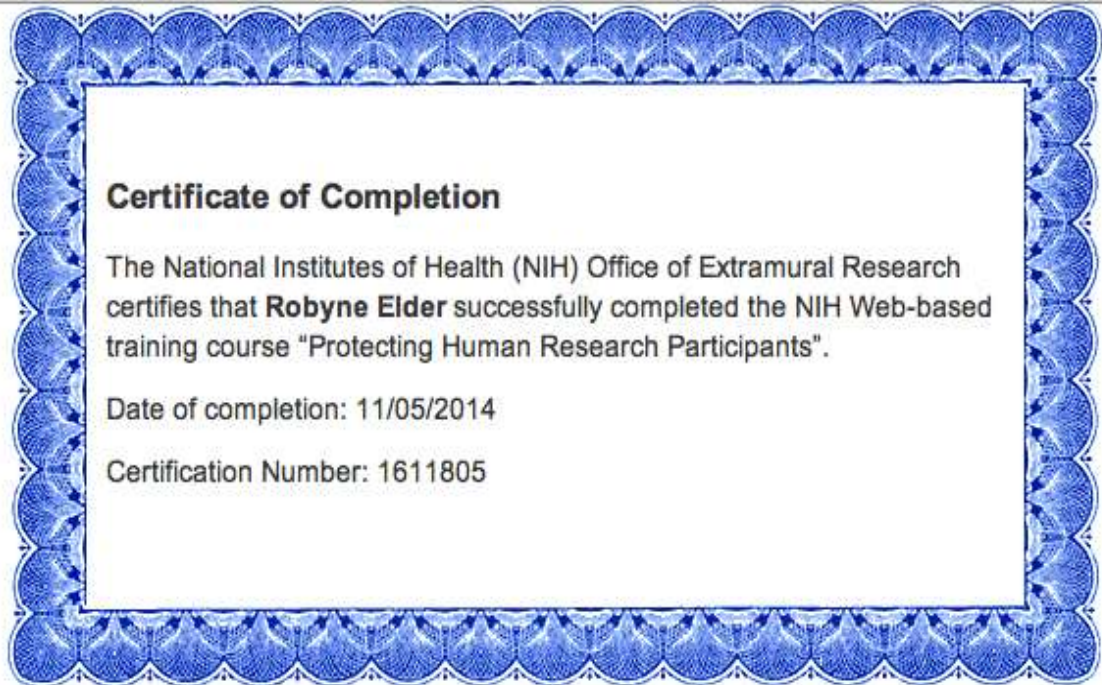
Negative behavior/feedback							Positive behavior/feedback		
1	2	3	4	5	6	7	8	9	10

Notes:

Scorers' behavior/feedback toward the teacher preparation program and students' readiness for the classroom:

Negative behavior/feedback							Positive behavior/feedback		
1	2	3	4	5	6	7	8	9	10

Notes:

Appendix D: NIH Certificate

Appendix E: Informed Consent

Thank you so much for participating in the MoPTA training in scoring sessions. As you know from my participation in the MoPTA training sessions, my name is Robyne Elder, and I am a doctoral candidate, graduate assistant, adjunct instructor, and APA editor at Lindenwood University. The research I wish to conduct for my doctoral dissertation (A Mixed-Method Investigation of the Missouri Pre-Service Teacher Assessment Pilot Program at a Private Midwestern University) involves using your anonymous survey responses from the MoPTA training and scoring. By signing this form you consent to allow me to use your responses, anonymously, in my dissertation. If you have questions please do not hesitate to contact me during the scoring session and/or via e-mail.

Thanks so much,

Robyne Elder
Graduate Assistant/Adjunct Instructor/APA Editor
Lindenwood University
relder@lindenwood.edu

Signature

Vitae

Colleges and Universities

1994-1998: Bachelor of Arts in English, emphasis in Creative Writing with a minor in History from the University of Missouri-Columbia; 1999-2000: Master of Arts in Teaching from Lindenwood University; 2013-present: pursuing Doctorate of Education in Instructional Leadership (expected graduation date in August of 2015) from Lindenwood University

Teaching and Editing Employment History

2009-present: APA Editor

2014-present: Copy Editor for the LU publication, *Journal of Educational Leadership in Action*

2013-present: Adjunct Instructor at Lindenwood University

2001-2013: English Language Arts Instructor and Department Chair for grades 9-12, at Ft. Zumwalt West High School

Spring of 2001: First Grade Instructor at Northview Elementary, Jennings School District

Awards

2011-2012: Educator of the Year for the Ft. Zumwalt School District

December 2011: Emerson Excellence in Teaching Award

2010-2011: Teacher of the Year for Ft. Zumwalt West High School

Spring 2010: Recognized as an Outstanding Teacher from the Honors College at the University of Missouri-Columbia