SEATTLE PACIFIC LIBRARY

Seattle Pacific University
Digital Commons @ SPU

Education Dissertations

Education, School of

Spring 5-11-2020

Video Analysis in Educator Preparation and Its Impact on Teacher Performance Assessment

Kirsten Koetje
Seattle Pacific University

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

 Part of the Teacher Education and Professional Development Commons

Recommended Citation

Koetje, Kirsten, "Video Analysis in Educator Preparation and Its Impact on Teacher Performance Assessment" (2020). *Education Dissertations*. 57.
https://digitalcommons.spu.edu/soe_etd/57

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

Video Analysis in Educator Preparation and Its Impact on Teacher Performance Assessment

by

KIRSTEN KOETJE

A dissertation submitted in partial fulfillment

Of the requirements for the degree of

Doctor of Philosophy in Education

Seattle Pacific University

May 2020

Seattle Pacific University

Video Analysis in Educator Preparation and Its Impact on Teacher Performance Assessment

By KIRSTEN KOETJE

A dissertation submitted in partial fulfillment

Of the requirements for the degree of

Doctor of Philosophy in Education

May 2020

Approved by

Dr. Nyaradzo Mvududu, EdD, Chair of the Dissertation Committee

Dr. Robin Henrikson, PhD, Committee Member

Dr. David Denton, EdD, Committee Member

Program Authorized to Offer Degree: School of Education

Date September 2020

Copyright Page

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

Signature 

Date 9/25/20

Acknowledgements

It is a strange thing to finish your dissertation amid a global pandemic. On the one hand, I am accomplishing a goal that I set mid-life that brings me much pride and satisfaction. I will celebrate the learning, commitment, and work that went into this effort. And yet, on the other hand, much is put in perspective when basic things like hugging your friends are no longer allowed. It feels almost trivial and self-centered to be expending energy here. Alas, I am not a medical doctor, nurse, or health care worker, so I toil on in my academic pursuits buoyed by the wisdom and encouragement of my chair, Dr. Nyaradzo Mvududu. I thank her for pulling me back into academic language, assisting me in the statistical analyses and verbiage, and lifting my spirit in this "stay at home" era via online Zumba sessions. Her calm and matter-of-fact demeanor has been a true gift to me in this season. Thank you Dr. Henrikson and Dr. Denton for giving me valuable feedback and for being general thought partners with me in educator preparation. I respect you both so much.

I am so grateful for my husband, Jon, and three children Jackson, Ashleigh, and Brooks. You have sacrificed many moments of being together to allow me to pursue this personal goal. Despite my full brain, you make me laugh and remind me of my deepest values. To my husband Jon, you supported me in very practical and emotional ways throughout this pursuit. I am so thankful for your consistent words of encouragement. I am proud of us for supporting each other's dreams!

Table of Contents

List of Tables iii

List of Figures iv

Chapter 1: Introduction 2

Theoretical Framework and Video Analysis 12

Chapter 2: Literature Review 17

Search Criteria 20

Who Should be in the Video—Others or Self? 22

Video Analysis with Pre-service Teachers 28

Meta-analyses Regarding Video Self-Analysis 31

Other Video Analysis Literature 38

Chapter 3: Methodology 41

Research Design 41

Research Questions and Hypothesis 42

Participants 43

edTPA Scores as Outcome Variable 52

Chapter 4: Results 56

Preliminary Data Analysis 56

Relationship between Writing Ability and edTPA Performance 59

Relationship between Video Analysis and edTPA Scores 60

Secondary Findings 61

Chapter 5: Discussion 64

Bias towards Strong Academic Writers in edTPA Unsupported 64

Critiquing Self Versus Similar Others.....67

Simulating a Form of Video Clubs—Critiquing Self and Other in Same Activity 70

Limitations.....72

Implications 75

Recommendations for Future Research.....79

Conclusion.....80

References.....82

List of Tables

Table 1. Summary of Selected Studies 22

Table 2. Group Means and Percentage of Maximum Critical Incidents (Seidel et al., 2011) 25

Table 3. Study Groupings and Omnibus Effect Sizes (Morin et al., 2019) 36

Table 4. Intern edTPA Endorsements 44

Table 5. Focus of Guiding Questions in edTPA Rubrics 53

Table 6. Descriptive Statistics of Variables 56

Table 7. Correlations of Predictor Variables to edTPA Performance 59

Table 8. Correlation Table of Disaggregated and Composite Variables 61

Table 9. Two Cohort Years' Means 63

List of Figures

Figure 1. Scatter Plot of “Number of Videos Critiqued” with Outlier 57

List of Appendices

Appendix A: Online Supervision Timeline 89

Appendix B: EPP Reflection Assignment and Rubric for Student Teaching 94

Appendix C: Video Coaching Platform Illustration of Markers and IPC Alignment.....96

Appendix D: Sample Internship Performance Criteria (IPC) Molar Rating Scale 97

Appendix E: Sample Video Analysis Comments 98

Abstract

The researcher analyzed data from two online cohorts of preservice teachers at a small, liberal arts university in Washington State. The researcher conducted a correlational analysis to determine if standardized writing scores and the quantity of video analysis conducted during the educator preparation program (EPP) predicted performance on the national Educative Teacher Performance Assessment (edTPA). Contrary to the researcher's hypothesis, academic writing ability did not have a predictive positive relationship with edTPA performance, $r = -.004$, $p = .98$. The total quantity of video analysis, both of *self* and of *other*, resulted in a statistically significant positive correlation with total edTPA scores, $r = .34$, $p = .02$. Disaggregated data showed significant positive correlations with the quantity of *others'* videos interns commented on, particularly with edTPA Task 3: *Assessment.*

Keywords: video analysis, preservice teacher, video feedback, teacher preparation, teacher self-efficacy, structured video analysis, video clubs, edTPA, standardized writing tests

Chapter 1: Introduction

Many scholars claim that of the variables over which school systems have some control, teacher quality and personal factors contribute the most significantly to student outcomes (Darling-Hammond, 2006b; Hattie, 2012; Marzano, 2007; Nagro, deBettencourt, Rosenberg, Carran, & Weiss, 2016). Educator preparation programs (EPPs) have faced fierce criticisms throughout the decades, accusations of not providing rigor nor professional skills to actual classroom teaching practice (Darling-Hammond, 2006a & b; Goldhaber, Cowan, & Theobold, 2017; Parkes & Powell, 2015). Teacher candidates themselves perennially critique EPPs for having theoretical learning in coursework that seems to diverge from the "reality" they face in the classroom (Schieble, Vetter, & Meacham, 2015). As one example, Riley (2020) explained that a group of five EPPs wanted to be more strategic in implementing reading science into their elementary educator programs. These EPPs reported that while a majority of candidates (60-80%) had opportunities to learn about reading science in coursework, a majority of those teachers (50-60%) received no practical training on actually employing the reading science principles in the classroom (Riley, 2020).

The most common connection between coursework to actual classrooms is student teaching internships. Often at the end of coursework, and sometimes simultaneously, teacher candidates complete teaching internships, which include live observations from mentors and field supervisors. These internships give candidates an opportunity to operationalize their learning. The live, face-to-face internship observations and debriefs that occur between teacher candidates and their supervisors are typically

limited to the memory of what transpired, from both perspectives. One promising practice bridges the gap between research and everyday classroom actualities: video analysis.

Analyzing one's own teaching in authentic classroom contexts via video recordings can connect the theoretical to the practical, a form of theory in action (Beck, King, & Marshall, 2002; Blomberg, Sherin, Renkl, Glogger, & Seidel, 2014; Marker & D'Onfrio, 2010; Nagro et al., 2016; Schieble et al., 2015; Seidel, Blomberg, & Renkl, 2013). For this paper, video analysis includes video recordings of actual classroom lessons for the purpose of instructional observation, review, and/or reflection. With video, teacher candidates have the literal time and opportunity to pause, rewind, watch without audio, watch with only audio, re-watch, or get a bird's eye view of the classroom. Most importantly, video allows candidates and supervisors to focus on a particular section of a lesson with a particular instructional focus without urgent moment-to-moment classroom demands or emotional charge. Relying on in-the-moment awareness or memory alone can be quite limiting and subjective, and even at odds with reality (Knight, 2014; Yusuf, 2006). When a teacher can watch the same clip through various lenses, such as with the focus of building classroom rapport, linking prior knowledge, student-to-student interactions, or distinguishing the types of verbal feedback given, different aspects may be noticed while reducing the cognitive demand for each viewing (Beck et al., 2002; Blomberg et al., 2013; Schieble et al, 2015).

Preparation programs typically aim to design activities that allow for deep reflection. The intent is that structured self-reflection will lead to positive instructional changes or affirmation to continue effective practices (Brownell et al., 2019; Mena-Marcos et al., 2013). Video analysis can lead to deeper reflection of one's efforts because

it allows for the multiplicity of access by self and others to the same scenario. Teacher educators want to foster this kind of deep reflection in teacher candidates so that these ways of thinking will transfer to their classroom behaviors (Beck et al., 2002; Beisiegel et al., 2018; Brownell et al., 2019; Dawson et al., 1975; Kimbrough et al., 2008; Nagro et al., 2016; Santagata & Sandholtz, 2019; Sherin & van Es, 2009). One question for teacher educators persists: What learning activities promote structured self-reflection that lead to instructional improvement?

While teacher educators aim to design effective, relational programs that develop reflective practitioners, they balance the need to offer their programs in ways that can reach and serve wider audiences with broad geographic expanse. In some locales and endorsement areas, teacher shortages grow more concentrated (Adams & Manuel, 2016). To meet the need of teacher shortage concerns, a proliferation of alternative route programs has been established to make pathways for adults who decide post-college to enter the teaching profession. Many of these programs have adopted a blended or online approach to teacher certification in order to offer flexibility and cast a wide geographic net for accessibility purposes since many rural and remote regions do not have an EPP. One challenge presented to EPPs involves developing online models that offer similar quality and support to candidates in a telecommuting environment as the traditional brick and mortar programs. Some research indicates that the type of teacher certification program—traditional or alternative--does not have as much impact on the teacher candidate as the actual practices in the program (Cole, 2018). Brownell et al. (2019) explain that novice classroom teachers often cling fast to rules and strategies, limited in their flexibility due to lack of experience. Integrating knowledge into situated contexts

involves flexibility, a skill expert teachers develop over time. Thus, EPPs of all models need to leverage highly effective practices for training new educators in their pedagogical skills and self-reflective practice to become efficacious teachers.

Reflection in teacher education remains a cornerstone practice for growth (Darling-Hammond, 2006a, 2006b; Mena-Marcos et al., 2013; Nagro et al., 2016; Schieble et al., 2015; Sherin & van Es, 2003). Self-reflective strategies promote effective educator actions and behaviors (Darling-Hammond 2006a, 2006b; Marker & D'Onfrio 2010). Video self-assessment provides a key component to having teacher candidates reflect upon their practice and consider the effectiveness of their choices. Although reflection remains a foundation in teacher preparation, reflection alone is not sufficient (Brownell, 2019; Mena-Marcos et al., 2013). A teacher candidate must go from being aware and reflecting on certain incidents to improving instructional practices through deliberate action (Mena-Marcos et al., 2013; Nagro et al., 2016; Schieble et al., 2015). Many states have mandated that pre-service teachers demonstrate this reflective skill along with their actual teaching aptitude in a performance portfolio. The national performance portfolio for pre-service teachers is called the educative Teacher Performance Assessment, better known as the edTPA. The edTPA involves self-reflective practice in authentic situations by mandating classroom artifacts, such as lesson plans, video clips, and student work samples with accompanying reflective analysis.

The edTPA is a standardized national performance assessment for those seeking teacher certification. The assessment purports to evaluate authentic evidence of teacher skills, rather than relying solely on concrete content knowledge (Greenblatt & O'Hara, 2015). The edTPA involves the teacher candidate creating a portfolio of artifacts and

commentary revolving around three classroom tasks: *Planning* (Task 1), *Instruction* (Task 2), and *Assessment* (Task 3). The artifacts that candidates submit to the external scoring agency include 3-5 consecutive lesson plans, lesson materials, video clips, assessments, scoring criteria or rubrics, and student work samples with teacher feedback. These artifacts come from the candidate's own student teaching classroom and thus provide evidence of the preservice teacher enacting their pedagogical knowledge. In addition to these artifacts, candidates provide 40-60 pages of analytical commentary on their instructional choices and implementation of the lessons. Santagata and Sandholtz (2019) pointed out that performance tests, such as the edTPA or the Performance Assessment of California Teachers (PACT), require preservice teachers to demonstrate their own teaching, rather than simply analyze someone else's teaching or prove content knowledge on an exam. In other words, the edTPA requires that teachers apply their pedagogical learning and content knowledge and demonstrate their ability to enact their learning for the purpose of actual, not hypothetical, student learning.

The edTPA stems from the PACT, which aimed at providing a standardized way to measure teaching skill without relying solely on the judgment of individual teacher educators or EPPs. The content of the edTPA was developed at Stanford University by the Stanford Center for Assessment, Learning and Equity (SCALE). A broad network of professional educators, including over 1,000 educators and more than 450 institutions of higher education, informed its development (American Association of Colleges of Teacher Education [AACTE], 2017). All three edTPA tasks include written commentary by the candidates asking them to justify their original decisions as well as reflect on the actual implementation of the planned lessons (Goldhaber et al., 2017; Parkes & Powell,

2015; Pecheone et al., 2016). With all of the written commentaries, explanation, and justification, the edTPA proves to be both an authentic teaching task while also a substantial writing task (Greenblatt & O'Hara, 2015; Santagata & Sandholtz, 2019). Some researchers contend that although teacher performance is the goal, the amount of reading, writing, and technical savvy required for the finished portfolio may conflate final scores alleging to measure teaching competence (Greenblatt & O'Hara, 2015; Kim, 2019; Santagata & Sandholtz, 2019). It is such a common conception among EPPs that edTPA relies upon strong academic writing that Whittaker, Pecheone, and Stansbury (2018) highlight this claim in their report responding to various edTPA critiques. They counter that scorers are trained to be aware of this potential bias regarding writing quality. Additionally, they state that as of the date of their rebuttal report in January 2018, no empirical evidence had been published to support this critique (Whittaker et al., 2018). They cited an unpublished study that in fact showed no correlation between standardized writing scores for program entrance and edTPA performance. Thus, this common conception of strong academic writing influencing edTPA scores continues to propagate without empirical support. Part of this current study intends to shed some light and provide evidence on this assertion.

There is some evidence to suggest that passing teacher performance assessments adds modest predictive value to future students' math and reading scores (Goldhaber et al., 2017; Newton, 2010). For instance, public school students who were assigned to a teacher who had passed the edTPA scored .252 standard deviations higher in reading than their counterparts who had been assigned to a teacher who failed the edTPA. The researchers controlled for other factors, such as which EPP a candidate completed and the

school district they taught in (Goldhaber et al., 2017). Newton (2010) conducted a value-added measure study of PACT scores to see if pre-service teacher performance could be linked to future student achievement on standardized tests. He controlled for student variables such as socioeconomic status and prior achievement to see if teachers who performed well on the PACT had a more positive effect on student learning trajectory than their lower performing teacher counterparts. Newton (2010) found that the main predictors of future student achievement, though modest, came from pre-service candidates scoring high in the PACT areas of a) assessment and b) descriptions of student language development (part of academic language rubrics in the edTPA). Value added measure studies, such as those of Newton (2010) and Goldhaber et al. (2017), which use teacher performance assessments to predict future K-12 student achievement, are typically limited to reading and math scores, since those are the domains of standardized testing so closely documented in our current educational system. Standardized assessment data in history, drama, world language, visual arts, music, and physical education do not exist at the same level for K-12 students, and thus, researchers generally do not design studies around them.

Proponents of the edTPA claim that it provides an authentic and predictive task of teaching practice that requires candidates to reflect on student learning (Goldhaber et al., 2017; Darling-Hammond, 2006a; Parkes & Powell, 2015). However, many question the validity and authenticity of a task that is highly edited and curated for the purpose of passing a performance test. For instance, since teacher candidates may choose to edit out instructional missteps, choose only their best lessons, select favorite subject matter, and even choose the students to highlight in video submissions and work samples, critics

challenge the claim of "authenticity" (Greenblatt & O'Hara, 2015). Promoters of edTPA policy advocate its real classroom context in K-12 classrooms as demonstrative of necessary teaching skills, rather than relying solely on pen and paper tests or contrived preparation program case studies or role plays (AACTE, 2017; Darling-Hammond, 2006a; Newton, 2010; Pecheone et al., 2016).

Teacher educators strive to prepare skilled teachers who can positively influence student learning and achievement. Teachers need to know their content knowledge, their students, and subject-specific pedagogies for facilitating their students' learning. However, going beyond knowing and completing coursework or a multiple-choice test, teacher candidates must demonstrate that they can enact this knowledge in an authentic classroom full of students. There is a difference between knowing or believing a certain action will have a particular outcome, known as an outcome expectation, and knowing how to motivate and control one's own behavior to enact that knowledge, or a self-efficacy expectation. EPPs have both tasks of presenting the various effective pedagogical strategies that lead to learning, and the challenge to help students operationalize this knowledge in situated classroom contexts (Brownell et al., 2019). Student teaching internships provide this opportunity, but they also demand a lot in terms of time and emotional investment. Pre-service teachers are simultaneously learning the school systems, mentor style, curriculum, and getting to know the students and families. Student teaching poses a large cognitive demand, and interns often feel overwhelmed. While live observations from mentors and supervisors can help candidates process instructional events, live observations have limitations. Ceven McNally (2015) reported a lack of structure, or the absence of specific protocols for classroom observations, as a

common difficulty. Oftentimes classroom observations end with disorderly notes without any implication for actionable follow-up goals. This lack of structure or focus can hinder both live and recorded observations. Video analysis allows for breaking up the viewing into smaller, more manageable parts, or what educators refer to as chunking (Marzano, 2007). A viewer can focus on the 5-minute opener of a lesson, or a 5-minute review and closure. It can focus on student-to-student engagement, or teacher-to-student positive to corrective ratio. Video analysis also offers a bridge to theory. It can take the complex practice of teaching and allow stakeholders to focus on one or two priorities at a time. By selecting a focus question, video provides a bridge based upon the declarative knowledge and research-based principles learned in EPP coursework to the contextualized classroom. What do the candidates notice about their own implementation? Students can self-reflect on their own practice, without simply relying on an outsider's feedback. Candidates get to observe themselves alongside their mentor and field supervisors. Candidates may be less defensive and more open to critical feedback if they notice something themselves (Knight, 2014). For instance, if a world language teacher aims to conduct 90% of her class in the target language, how much class time is actually devoted to the target language versus English? Video analysis offers the opportunity for the candidate herself to collect this data. Self-reflection of these classroom incidents can also provide powerful motivation for adult learners to determine their future goals. Andragogy literature suggests general principles of adult learning theory, namely that adult learning is most effective when it is self-directed, goal-oriented, relevant, voluntary, and often practical in nature (Zepeda, 2012).

Problem Statement

Teacher educators aim to prepare reflective and analytical practitioners who can adjust their own teaching for the sake of student learning. The literature suggests that video analysis offers one tool for developing self-reflective practitioners. The current measure used in Washington State, and elsewhere in the nation, for evaluating pre-service teacher competence is the edTPA. The purpose of this study is to explore if the experience with video analysis of authentic classroom teaching during the preparation program positively impacts teacher candidates' edTPA scores, or their demonstration of teaching competence. Due to the large writing demand of the edTPA, producing around 40-60 pages of writing, students who come in with strong writing skills may have an advantage when creating their portfolio (Greenblatt & O'Hara, 2015; Santagata & Sandholtz, 2019). The researcher hypothesizes that there will be a positive correlation between program entrance writing scores and edTPA scores. In addition to strong writing skills, the researcher hypothesizes that the quantity of structured video analysis conducted by teacher candidates during the program will have a positive relationship with edTPA scores. The literature has shown that video self-analysis provides a powerful, active learning tool to help teacher candidates transfer learning from university coursework into their K-12 internship classrooms (Morin et al., 2019; Nagro et al., 2016). A growing body of research on video analysis suggests that it has the power to influence teacher behaviors and instructional change (Morin et al., 2019). However, the researcher has found a gap in the literature showing a relationship between video analysis efforts with candidates' performance on the edTPA.

Recording lessons of authentic classroom interactions allows candidates to connect their learning to their actual daily practice and identify evidence, or lack thereof,

of their growing pedagogical skill. Video self-analysis can complement research, coursework, and workshops by giving candidates the ability to target a particular practice, implement it in their regular classroom, and reflect on the implementation's effectiveness, thereby providing legitimate follow-up opportunities while transitioning from knowledge acquisition to application (Knight, 2014; Morin et al., 2019). Video analysis of one's own teaching has been used in a variety of educational settings (e.g., one-on-one instruction and whole class), grade levels, and for both pre-service and in-service teachers (Morin et al., 2019), and also in a multitude of interactional professions (Fukkink et al., 2011).

Theoretical Framework and Video Analysis

Bandura's (1994) theory of self-efficacy, a facet of his social cognitive theory, posited the notion that nurturing individuals' beliefs about themselves to enact certain behaviors can impact their motivation. His social cognitive theory frames a reciprocal relationship between three components: one's behavior, internal personal factors, and the environment (Bandura, 1982). EPPs have little influence on candidates' future school environments and internal personal factors, such as genetic inheritance or personality. Accordingly, EPPs emphasize teacher behaviors and ways of thinking. Guided video analysis offers a tool for focusing reflective thinking that can influence behaviors (Beck et al., 2002; Blomberg et al., 2014; Knight, 2014; Morin et al., 2019; Nagro et al., 2016; Schieble et al., 2015; Seidel et al., 2013; Sherin & van Es, 2003; Sherin & van Es, 2009).

Bandura noticed the often-incompatible nature between human knowledge and human behaviors. To summarize this phenomenon of contrasting knowledge and consequent action, Bandura (1982) concluded that efficacious persons have a sense of

being able to influence and regulate events around them. Bandura (1994) theorized many benefits of highly self-efficacious individuals, such as approaching challenges with a goal orientation, whereas he warned of significant liabilities for those less self-efficacious, such as avoiding difficult tasks and perceiving them as personal threats. Video analysis offers coherence between knowledge and teachers' own daily instructional practice in context, so that teachers may form specific goals. If the goal is 90% target language immersion, that goal may be applied and then reflected upon for meaningful and long-term follow-up in future instructional recordings (Morin et al., 2019).

In a classroom, there are many variables outside of a teacher's control, such as classroom assignment, students assigned, mandatory testing, and curriculum. However, numerous other factors lie within teachers' decision-making powers—for example, day-to-day planning, student rapport, classroom setup, interactions with students, engagement with families, classroom management, professional relationships, general attitude, attire, and so on. These quotidian classroom activities reside within the teacher's control. Self-efficacious teachers set goals and maintain a strong commitment to their instructional quality, despite challenges. Video self-analysis allows teachers to focus on a targeted practice within their control, and reflect upon their enactment of that practice without simultaneously having to balance all of the demands of teaching (Derry et al., 2010; Morin et al., 2019).

Bandura (1994) outlined four key areas that impact personal efficacy: 1) mastery experiences or successes, 2) vicarious experiences watching others similar to oneself modeling success (the higher the perceived similarity, the more persuasive), 3) social persuasion by verbal encouragement, and 4) one's own somatic and emotional state with

perceptions of stress or pain being more influential than actual levels. Of the four, Bandura (1994) claimed personal mastery successes had the greatest impact. In other words, people who experience first-hand positive achievements are more motivated to replicate their own behaviors. According to Bandura, in order to instill resilience, achievements that are somewhat hard-won have more impact than a series of easy successes. For social persuasion, he termed knowledgeable encouragers as *efficacy builders*. He noted that efficacy builders put others in situations where they can experience reasonable growth, rather than prematurely assigning them to extremely challenging situations. Efficacy builders prioritize self-improvement over competitive triumphs (Bandura, 1994). Thus, the most influential mastery experiences seem to need some amount of obstacle overcoming, but not too much, with encouragement from a knowledgeable outsider. Bandura (1994) warned that social persuasion could also easily undermine personal self-efficacy—verbal discouragement or attention to weaknesses—rather than nurture it. For social persuasion to have a lasting impact, he advocated realistic affirmations paired with noticeable successes. Applying Bandura's (1994) theory of self-efficacy to video analysis, teacher candidates can view/notice their own mastery successes, benefit from social persuasion in the form of positive encouragement from mentors, peers, and supervisors; and reflect on their own ability to take on classroom challenges.

When trying to build self-efficacy, the difference between an outcome expectation and an efficacy expectation becomes a critical distinction. Bandura (1977) delineated the distinction between the two explaining that an outcome expectation describes what people believe will likely happen if they perform a certain action, while an efficacy

expectation describes one's own conviction regarding the personal execution of said action. This difference has great significance because people can estimate that a particular behavior will lead to positive outcomes, but doubt their own ability or motivation to perform it. Thus, what teachers know or believe, or their particular outcome expectations regarding certain teacher behaviors, will not necessarily translate into what teachers believe they can or will do, self-efficacy expectations. For instance, one may have the head knowledge that learning and correctly using 200 student names will have positive outcomes, but the self-discipline and motivation to enact that behavior does not automatically follow. Video self-analysis offers one potential strategy for providing this motivation and self-belief in two helpful ways. If reflective practitioners notice a success, they can highlight this internally and reproduce more of the positive behaviors. Effective feedback consists of identifying desirable performance to encourage replication of that behavior (Hattie & Timperley, 2007). If reflective candidates notice something to improve, they can set a specific and relevant goal (e.g., increase from 60% target language to 80% target language use), and use video for on-going development and follow-up to make observable progress towards that goal (Morin et al., 2019; Zepeda, 2012).

In the particular case of this study, the video analysis used in the EPP revolved around two major anchor rubrics, the Internship Performance Criteria (IPC) and the edTPA rubric language. The IPC is the culminating assessment for the university's yearlong internship. Interns refer to the IPC throughout the year to discuss growth. The IPC includes eight main criteria which align to Washington State teaching standards. Teacher candidates being supervised online had to record lesson observations for their

supervisors according to the timeline set out in Appendix A. Candidates could then select which lesson recordings to watch and use the instructional recordings as evidence for self-reflection. Candidates were asked to reflect on their classroom practice by focusing on one of eight IPC criteria. Interns may have self-selected the criteria or agreed upon the criteria after conversing with their mentor or university supervisor. In the reflections, interns described a classroom situation or scenario and then reflected on its implementation and suggested possible means of improvement. Interns are encouraged to set goals in their written reflections. See Appendix B for the specific instructions for these intern reflections. In addition to the IPC-based written reflections, teacher candidates also used the video coaching platform to timestamp selected videos when they saw evidence of a particular rubric. Interns and supervisors can add rubric markers and comments in the coaching platform that align to either the IPC or that use edTPA rubric language, such as the development of K-12 students' academic language. The IPC and edTPA provide reflective prompts to focus the interns' attention. See Appendix D for an example of *Expectations,* one of the university's eight IPC criteria, and the accompanying molar rating scales interns could have chosen to focus their viewing. After selecting a reflective focus, the interns can then set a specific goal. Andragogy literature suggests that adults are more motivated when they can self-select relevant and practical goals for their own development (Zepeda, 2012). Self-reflection and goal-setting can in turn provide a pathway for achieving mastery success, which has the potential to build candidates' internal self-efficacy (Bandura, 1994).

Chapter 2: Literature Review

For the past 100 years or more, teacher preparation in the United States has been surrounded by controversy and even debate over its very existence (Bohan, 2016; Darling-Hammond, 2006a, 2006b; Kennedy, 2015). Questions abound as to what makes the most effective teacher preparation experience, which stems from the question, "What makes the most effective teacher?" Since most Americans have had personal experience in the K-12 classroom, many want to contribute to the public education conversation with personal anecdotes and experiences. Teacher preparation has been critiqued for not adding value to what the common citizen already knows, having non-prepared citizens score nearly as well on teacher preparation assessments as those who complete preparation programs (Darling-Hammond, 2006a). EPPs face the valid critique of over-relying on theory and university coursework, at the expense of devoting time to actual classroom implementation (Riley, 2020). Modern learning theory challenges our understanding of what it means to learn from a more simplistic transmission model of information to a more complex notion of varying learner backgrounds, varying goals, and the ability to transfer and apply knowledge and processes to various contexts (National Academies of Science, Engineering, and Medicine [NASEM], 2018).

The landscape of educator preparation continues to undergo reform as new research appears and policies change. Whereas teaching used to be seen as a more routinized process which could be reduced to a list of pre-determined activities to become a proficient "technician," teacher educators now prioritize the role of teacher as reflective and responsive diagnostician, knowing how to effectively address challenges inherent within the profession depending upon context (Darling-Hammond, 2006b; Kennedy,

2015). Thus, it remains in the forefront of current EPPs to equip teachers for the more demanding role of facilitating deep learning, firmly grounded in modern learning theory. Simultaneously, EPP programs strive to nurture teacher self-efficacy amidst a myriad of on-going classroom challenges. At the vanguard of this national push to view teachers as responsive and reflective practitioners, rather than skilled technicians, comes the relatively nascent Educative Teacher Performance Assessment (edTPA). The edTPA is used as a means of evaluating classroom evidence that candidates can translate knowledge into practical, meaningful learning experiences for their specific students. Creators of the edTPA intended for the performance assessment to be a more authentic measure of *teacher as reflective practitioner* - one that measured actual learning facilitation, rather than merely measuring discrete knowledge on a test. Darling-Hammond (2006b) enumerates the various kinds of active learning experiences that exemplar EPPs utilize to prepare their teacher candidates to become responsive educators. She highlights activities such as child observation case studies, micro-teaching, analysis of authentic classroom artifacts (e.g., students' work samples, video recordings), and portfolios. Videos of classroom instruction have become increasingly available and convenient with modern cell phone and computer technology. The edTPA provides an opportunity for teacher candidates to create a portfolio of their actual classroom context, not a generalized one, and adhere to coherent national teaching standards and principles of quality instruction. With active learning strategies and teacher portfolios at the center of teacher educators' current conversations, what role does the growing use of video analysis have in EPPs? Can video analysis facilitate deeper learning in teacher candidates and influence personal instructional change?

Video self-analysis provides a suitable strategy for pairing the theoretical with the practical during teacher preparation; it offers a bridge (Beck et al., 2002; Blomberg et al., 2013; Blomberg et al., 2014; Marker & D'Onfrio, 2010). With the ubiquity of smart phones and recording devices, instructional video has become more and more common and affordable (Knight, 2014; Morin et al., 2019). Both the edTPA for pre-service and National Board Certification for veteran teachers require a video analysis component, indicating video's significant place within the teaching profession. Using the authentic artifact of personal classroom recordings, video analysis can provide pre-service teachers a vehicle for really seeing their practice rather than relying simply on what they believe happened (Knight, 2014). Video analysis of a teacher's own practice offers the potential for deeper teacher reflection and greater motivation (Beisiegel et al., 2018; Knight, 2014; Nagro et al., 2016; Seidel et al., 2011). In Beisiegel et al.'s (2018) study, math teachers in a professional development group who watched their own videos and were led by a teacher member slightly outperformed other group conditions who were led by trained facilitators and/or watched stock video examples. Beisiegel et al.'s (2018) study used an outcome measure called the Mathematical Quality of Instruction (MQI), which is a video observation tool used to analyze elementary math instruction. Teachers' perceptions and beliefs of their own ability to be successful in a classroom context, or their teacher self-efficacy, can influence their actual practices and motivation. Video analysis allows a teacher to notice salient aspects of teaching interactions and to focus on student learning (Beisiegel et al., 2018; Santagata & Sandholtz, 2019; Sherin & van Es, 2009). In addition to watching their own videos, teachers reported that peers often noticed learning moments that they did not and that they learned by watching their peers' instructional

maneuvers (Beisiegel et al., 2018). Thus, encouraging and empowering teacher self-efficacy through guided video analysis and feedback provides one practical strategy EPPs can use to impact teacher behaviors and habits of mind, with the hope of in turn positively influencing student learning.

Search Criteria

When researching literature on video analysis in teacher preparation programs, the researcher used the following search terms in Academic Search Complete: video analysis, video self-analysis, self-efficacy, teacher preparation, educator preparation, pre-service teacher, and teacher candidates. Preference was given to studies designed with pre-service teachers since that is the context of the research study at hand, but studies including practicing teachers were also included. Studies were selected if they involved teacher self-analysis of their own authentic teaching in the K-12 environment or if there was video analysis as an independent variable, with the outcome measure involving instructional performance.

For the purposes of this study, during the literature review, the researcher prioritized articles that focused on video analysis in teacher development, preferably teacher candidates, with the inclusion of a quantifiable external evaluation of actual teacher performance. In the search results, there were varying video analysis foci, including discipline-specific such as video analysis for math or reading or the difference between watching one's own video versus another's instruction. Table 1 summarizes the contributions of five selected articles germane to this present study's focus on video analysis influencing pre-service teachers' skills in actual K-12 classrooms. Table 1 shows that no other selected study from this literature review combines all of the areas of focus

in this present study, which includes an externally scored observation of teacher skill as the outcome variable, edTPA in this case. None of the other selected studies attempted to factor in previous writing or academic ability, and only two others focused exclusively on the development of pre-service teachers. With some of the literature suggesting that students with stronger academic backgrounds, regardless of actual teacher competence, will have an advantage, this study attempts to factor in writing skill by using Washington Educator Skills Test – Basic for Writing (WEST-B Writing) scores as a proxy (Greenblatt & O'Hara, 2015; Santagata & Sandholtz, 2019). The author wants to examine if the quantity of video analysis will predict success and teacher competence as demonstrated in the edTPA, without conflating that success with writing ability. With this dissertation, informed by these other scholars' work, the author's goal is to provide some information to fill the gap on the use of video self-analysis in EPPs to develop pre-service educators' teaching competencies as evaluated by external scorers using coherent national standards.

Table 1

Summary of Selected Studies

Study	Externally scored	Writing ability	Focus on self-analysis	Comparing analysis of self- vs. other video	Focus on pre-service teachers
Present study	X	X	X		X
Nagro et al. (2016)	--		X		X
Santagata & Sandholtz (2019)	X				X
Fukkink et al. (2011)	--		X		
Seidel et al. (2011)				X	
Morin et al. (2019)	--		X		

Note. The dashes stand for the presence of an evaluation by a scorer, not self-report, but the score was internally completed rather than by an external scoring agency, such as the edTPA or PACT.

Who Should be in the Video—Others or Self?

In the literature review, a couple of studies concentrated on the subject in the video. In essence, the researchers wanted to know if there was a difference in watching stock footage versus watching one's own classroom teaching (Beisiegel et al., 2018; Seidel et al., 2011). The general conclusion was that teachers preferred (self-reporting) watching their own video tape (Seidel et al., 2011), performed better when participating in professional development groups that focused on their own classroom video (Beisiegel et al., 2018), but were less critical with their own instruction or analyzing those of peers (Beisiegel et al., 2018; Seidel et al., 2011).

Since video has become quite common in the development of educators, Seidel et al. (2011) examined if the person in the video made a difference. As teacher educators, the researchers wondered if watching one's own video versus stock video made a difference. Since group dynamics can evolve over time and involve other social dynamics, they designed a quasi-experimental study that focused on individual video

analysis rather than analyzing tape as a group experience, such as in class or video clubs. Seidel and team (2011) also wondered if previous work with video tape would impact their analyses. This study fit under a 6-year project investigating science-teaching practices in Germany and Switzerland. From a previous stage in the study, they taped German physics teachers' lessons and thus already had a bank of physics classroom recordings as well as a pool of teachers experienced in video analysis. From the previous study, researchers identified video lessons of mechanics that offered a representative exhibition of mechanics instruction throughout Germany.

One of the main points of inquiry in the Seidel et al. (2011) study was to examine the effect of video material on video analysis as an activating experience. As an activating experience, the researchers wanted to know what kind of effect prior experience with video and the category of video (one's own vs. another's) would have on teachers' levels of *immersion* (deep level engagement), *resonance* (link to own teaching), *authenticity*, and *motivation*. The study involved the comparison among three groups of physics teachers ($N = 67$) who all watched a 45-minute lesson on introductory mechanics: 1) teachers who had experience with video analysis who watched their own lesson, 2) teachers who had experience with video who watched another teacher's lesson, and 3) teachers with no known video analysis experience who watched another teacher's lesson. All participants had taught the introductory mechanics lesson as part of their own regular curriculum.

The Seidel et al. (2011) study adds further support that video analysis—whether of self or others—does have positive impacts on teacher learning and motivation. All three groups combined showed that a majority (65.7%) experienced a high sense of

immersion in the video activities, either “mostly” or “always” feeling “inside the lesson.” The “own” condition felt a particular *resonance* with the video experience, with 81.9% reporting “mostly” or “always” having their own instruction in mind while watching the video. Additionally, the entire sample’s majority found video tasks motivating (63.8%), and offered an authentic representation of science teaching (91.2%). Table 2 summarizes the three groups’ means for their self-reports of video analysis as an activating experience in the previously described categories: *immersion, resonance, authenticity,* and *motivating*. The teachers who watched their own teaching videos experienced higher activation as measured by self-reports of *immersion, resonance,* and *motivation*. They found no significant differences between groups in measures of *authenticity* since all three groups found the videos highly authentic. For *immersion,* group condition explained 25% of the variance. For the *resonance* variable, group membership explained 12% of the variance. Of note, the “own” group who watched their own instructional videos were the least likely to reflect on critical incidents noticed in the playback.

Table 2

Group Means and Percentage of Maximum Critical Incidents (Seidel et al., 2011)

	Own video (n = 23)	Video-experienced other (n = 15)	Video-inexperienced other (n = 29)
Immersion[1]	**2.22**	1.74	**1.45**
Resonance[1]	**2.64**	2.13	**2.00**
Authenticity[2]	2.25	2.29	2.04
Motivating[2]	**1.91**	1.80	**1.41**
Articulating critical incidents[3]	**20.2**	25.8	**29.2**

Note. Means printed in bold indicated statistical significance between the groups (post hoc).
[1] Teachers' reactions directly after commenting on video
[2]Teachers' evaluations within 14 days of data collection
[3]Indicates the percentage of the group that reached the highest level of critical reflection.

Seidel and team's (2011) study of physics teachers used experienced teachers already in the field; it was not a study specific to developing pre-service teachers. It also used self-reporting questionnaires as an outcome variable rather than an external score of observable instructional skill. Self-reporting data can be unreliable as it may be threatened by self-reporting biases and typically has larger measurement error (Field, 2013; Gall et al., 2007). While professional development designers, including teacher educators, aim to create activities that are motivating and authentic, the true aim is instructional improvement. This study provided information on teachers' reception of video analysis, but it did not offer evidence on video analysis as a tool for instructional change or skill improvement. **External (Non-insider) Scoring of Teacher Competency**

In order to examine the effect of video analysis on instructional change, it is appropriate to include an external evaluation of teacher performance as an outcome measure. In the literature review, several other studies used an evaluation of teacher

performance that did not rely solely on self-report (Fukkink et al., 2011; Morin et al., 2019; Nagro et al., 2016; Santagata & Sandholtz, 2019). However, only Santagata and Sandholtz (2019) used an external measure, not dependent upon the scores of those involved in the research study or familiar with the candidates. Santagata and Sandholtz (2019), teacher educators in California, wanted to know if there was a relationship between preservice teachers' performance on a teaching performance assessment and another measure of their competency analyzing students' mathematical thinking via video recordings. The study used data from 89 preservice elementary teachers enrolled in a one-year post-baccalaureate preparation program. The study used a correlational design using students' scores on the Performance Assessment for California Teachers (PACT), similar to the national edTPA, and a classroom video analysis (CVA) instrument used specifically for analyzing authentic classroom math lessons. Santagata and Sandholtz (2019) reported that both assessments had shown connections to teacher competence and student learning, though the CVA instrument had been used with practicing teachers rather than preservice teachers.

Preservice participants completed the CVA assessment by watching and analyzing ten elementary math video clips of real classroom lessons, not their own. For each video, they answered the prompt: "Discuss how the teacher and the student(s) in the clip interacted around the mathematical content." This took approximately one hour of their time and resulted in a final written analysis, which was scored by trained scorers. The CVA consisted of 10 short videos ranging from 1 to 3 minutes of real math interactions in a classroom and took approximately 1 hour to complete. Viewers analyzed the clips without access to the rubric since part of the measure was to see what the

participants would notice and how they would explain their rationale. Their analyses were scored for specificity among four CVA components—*Mathematical Content* (CVA-MC), *Suggestions for Improvement* (CVA-SI), *Student Thinking* (CVA-ST), and *Depth of Interpretation* (CVA-DOI).

The researchers found that those who scored high on the CVA, analyzing the mathematical thinking of students in stock footage, were not the same as the high performers on the PACT. The authors posited several potential explanations for the discrepancies in high performers between the two assessments, which they found puzzling since both measures have previous research supporting their usefulness in predicting math teacher skill. They noted that the PACT took roughly 8-10 weeks of concerted effort, and students knew the consequential outcomes, whereas the CVA took only 1 hour of their time and had no certification consequence. Thus, there may have been a difference in the effort put forth by individuals. Additionally, the PACT could be seen to advantage strong readers and writers, and may not be measuring solely math teaching competency, but confounding attributes such as technology competency, writing ability, conscientiousness, or SES of the students in the classroom. Some candidates may have received low scores due to sub-optimal writing skills or inattention to a particular requirement. Furthermore, the CVA required a critical lens applied to others' teaching, whereas the PACT required self-analysis. Evaluating others may allow for more freedom to critique (Beisiegel et al., 2018; Seidel et al., 2011), but adds the challenge of trying to understand the specific context of the lesson and the student. Lastly, due to the high stakes of the PACT, some candidates might have altered or even falsified documents in order to receive higher scores. The authors noted that other studies had reported

discrepancies between teacher performance assessments and supervisor/mentor judgments, so incongruities between measures are not novel to teacher assessment. Santagata and Sandholtz (2019) cited one study that showed significant associations with undergraduate grade point averages (GPAs) and grades in methods courses as positively correlating with performance on the PACT.

Santagata and Sandholtz (2019) used an externally scored outcome variable, PACT, as this present study aims to do using the national edTPA, rather than internal scores from the EPP or self-reports from the candidates themselves. However, the Santagata and Sandholtz correlational design used CVA scores, which relied upon one-time stock video analysis of mathematical thinking, and not candidates' own classroom contexts over time. The present study also uses a correlational design, but uses the quantity of self-analysis of candidates' own and peers' teaching contexts as the predictor variable. Moreover, the present study includes writing scores to examine if the frequency of video analysis explains unique variance over and above candidates' writing abilities, which is not something that Santagata and Sandholtz (2019) attempted to tease out.

Video Analysis with Pre-service Teachers

Like Santagata and Sandholtz (2019), Nagro et al. (2016) focused on video analysis specifically in training pre-service teachers. Nagro et al. (2016) used both self-report surveys and an outside evaluation of teacher competence, but the scorer was the main researcher who worked with the teacher candidates in their program, thus introducing potential scoring bias. Nagro et al. (2016) found evidence to support that more structured video analysis paired with advisor feedback had a more positive effect on teacher candidates' reflective abilities and teaching performance than less structured

analysis without feedback. The quasi-experimental study involved 36 teacher candidates in their student teaching internship. Participants were divided into two groups of video self-analysis. Both groups recorded their own classroom teaching a total of four times—beginning, early midpoint, late midpoint, and end of the internship. Both groups were asked to watch each video within 48 hours and then submit a written reflection using timestamp evidence and language from of an instructional skills rubric. The comparison group (n = 19) ended with the confirmation of their submissions. The treatment group (n = 17), labeled as the guided video analysis set, had two additional components to the video analysis task. In addition to the video analysis and written reflection, the students in the treatment group evaluated their own written reflections based on a rubric and submitted their self-ratings. After submission, the candidates in the treatment group also received written feedback from the lead author of the study, Nagro. Nagro's feedback pointed out exemplars from the candidates' submissions, referenced Danielson's instructional framework, and asked probing questions, but refrained from being too evaluative. Participants did not respond formally to the feedback. In sum, the treatment condition included a self-evaluation of the written reflection and they received external written feedback, neither of which was present in the comparison group.

Nagro et al. (2016) posed three questions for their study comparing two groups of pre-service teacher groups—1) guided video analysis with feedback, and 2) less guided video analysis. The study used three measures for group comparisons, a) a teacher candidate questionnaire about perceived teaching ability (self-report), b) an externally scored written reflection rubric, and c) an externally scored instructional skills rubric. Teacher candidates took the questionnaire pre- and post-intervention on their own

perceived teaching abilities on 13 items, such as overall teaching, communicating expectations for learning, explaining content, and using discussion techniques. The composite score became the *perceived ability* score. The lead author scored the first and last (fourth) written reflections submitted for both the comparison and treatment groups. She also scored the first and last (fourth) video submissions of each participant in both groups using the instructional skills rubric (n=15 in treatment, n=13 in comparison). The instructional skills rubric included six components within Domain 3 *Instruction* of Danielson's (2013) framework with four levels of proficiency.

Evidence did not support the first research question that guided video analysis would impact *perceived abilities*. While they found no between-group effect on candidates' own self-reporting, the mixed-model ANOVA reported a significant within-group difference across time, $F(1, 34) = 35.32$, $p < .001$. This suggests that the overall experience of both groups during student teaching, each having some form of compulsory video analysis, did positively influence *perceived ability*. The second research question regarding *reflective abilities* produced the largest effect between conditions and a significant interaction between group condition and time, $F(1, 34) = 33.09$, $p < .001$. The third research question investigated actual *instructional skills* based upon an outside scorer's evaluation. Results showed a significant within-group difference over time, $F(1, 26) = 16.76$, $p < .001$, as well as a significant interaction between group condition and time $F(1, 26) = 6.83$, $p < .01$. Hence, group conditions did have a significant impact for the variable *instructional skill,* though not as large as the *reflective ability* variable. The two groups' mean comparisons in *instructional skills* help illustrate the amount of improvement for each condition. The comparison group went from 37.14 to 40.17,

whereas the treatment group went from 35.04 in instructional skills to 47.78. In sum, the results of this study showed *perceived ability* went up for both groups over time using video analysis, but group conditions did not have a significant impact. Providing a structured reflective activity with the video analysis, paired with advisor feedback, produced a statistically significant difference on actual *instructional abilities* as judged by outside evaluators and an even larger effect on candidates' *reflective writing*.

One weakness of this study was the lack of neutral or externally trained scorers outside of the EPP. While this study did include some inter-rater reliability of student teachers' *instructional abilities* as evaluated by the submitted videos, they were still scored by internal faculty who knew the candidates and the study. Using edTPA scores removes this level of potential personal bias of knowing the candidates. Additionally, this study did not compare to candidates who had not completed any video analysis. It is possible that time in student teaching, without any video analysis, would improve candidates' *perceived ability* or *instructional abilities* in a similar fashion.

Meta-analyses Regarding Video Self-Analysis

Two studies in the literature review were meta-analyses that shed light on the format of video self-analysis. One focused on helping professions more broadly, while another focused solely on special education educators at various stages in their career. Fukkink, Trienekens, and Kramer (2011), researchers in the Netherlands, investigated the impact of video feedback (VF) on various professions, seeing the medium as a way to study the behavior of professionals in-depth, and also to investigate the influence of VF on those behaviors. They noted VF as a common training technique in various interpersonal professions such as teaching, counseling, and medicine. Many professional

fields are turning to video analysis as a helpful training tool. The researchers emphasized positive self-modeling as a means of boosting self-efficacy, which according to Bandura's social cognitive learning theory, leads to repetition of that behavior (Bandura, 1977). They conducted a literature review of experimental studies done on VF where participants watched videos of themselves, not others. They searched databases for certain key terms such as: video*, self-model*, and self-confrontation. Studies had to involve external evaluation using an observation instrument, rather than self-evaluation alone, and the studies had to have sufficient quantitative data to calculate an effect measure. While the observation measures were not self-evaluated, there still may have been internal scoring biases in many of the studies. In total, they included 33 studies in the meta-analysis.

Within their 33 studies, they found that on average, VF interventions lasted 10 weeks with 4.4 sessions (SD = 2.30). Participants were filmed for an average of 20 minutes per session. Recorded sessions were reviewed on average one week later. The majority of outcome measures were positive in nature (88%), such as active listening or authenticity, while some studies included negative outcome measures (e.g. nervousness or passivity). Overall, the meta-analysis revealed the VF interventions to have an aggregate medium effect size of 0.40, which was statistically significant ($p < .05$). This meets Hattie's (2012) recommended $d = 0.4$ "hinge point" for an above average intervention and adds support that video self-analysis is an effective tool for impacting teacher behavior change. In the Fukkink et al. (2011) meta-analysis, a few items stood out as having practical implications for how to design VF. The effect size was larger for positive outcome measures (0.41) compared to negative measures (0.28), and larger for

outcomes of molar skills (0.52) compared with microskills (0.32). They defined a microskill as a highly specific skill measured by tallying (e.g., number of questions asked) and a molar skill as something more broadly assessed on a rating scale (e.g., responsiveness). Additionally, using a structured observation form for the targeted skills produced the largest effect size (0.55) compared to not using a form (0.21). In summary, the combined regression model predicted the largest experimental effect for VF programs that included a standard observation form that measured positive, molar outcome measures. With these variables, the model predicted an effect size of 0.68 and explained 48% of the variance between studies. Accentuating the positive does appear to have a constructive influence on candidates' behavior, lending support to Bandura's theory of mastery successes and social persuasion building a person's self-efficacy. It could be that as evaluators affirm the desired behavior, they see it replicated, reinforcing a positive self-modeling approach. Nevertheless, accentuating the positive did not extinguish negative behaviors. Fukkink et al. (2011) proposed more experimental research be conducted specifically targeting both positive and negative feedback in VF to systematically compare the two.

In Fukkink et al.'s (2011) meta-analysis, they screened for studies that included a standard evaluation form. Structured forms allow participants to narrow in on the target behavior and focus their attention on specific aspects, thereby reducing cognitive processing demands. Nagro et al.'s (2016) study also included a structured form, and that is one of the general recommendations for implementing effective video analysis (Derry et al., 2010; Knight, 2014). In the Fukkink et al.'s (2011) meta-analysis, the explicit use of an evaluation form had evidence of a positive effect. This implies that structured focus

and feedback are critical for candidates developing professional skills. Video analysis experiments that included molar outcome measures on a scale, such as *rapport*, produced more of a positive effect than measuring specific microskills. This seems paradoxical considering that focus and specificity via an evaluation form enhanced VF interventions. However, the researchers posited two possible explanations. Firstly, counting specific microskills may reach a ceiling effect. For instance, once a certain number of open-ended questions has been asked, further instances may not improve the quality of interaction. Secondly, Fukkink et al. (2011) suggested that improvements in professional interactions involve qualitative aspects typically measured by molar rating scales, such as offering respect.

In conclusion, this meta-analysis on helping professions yielded some very practical implications for the use of VF in training programs to improve interactional behaviors. In particular, a structured observation form did indeed show statistically significant larger effects (ES = 0.55) over those VF interventions that had no such form (ES = 0.21). It should be noted that both forms of VF interventions, with and without the structured form, produced a positive effect size. As a meta-analysis, this study did not have its own intervention to replicate. It also veered outside of the realm of pre-service teachers; it included other professions and novice and experienced practitioners alike. There is a lack of generalizability to pre-service teacher training. Despite these limitations, the overall conclusions of this meta-analysis informed the design of the current study's use of video self-analysis using a Danielson-inspired framework with a molar rating scale, with all of the outcome items positive, rather than negative, in nature.

Morin et al. (2019) acknowledged the recent growing body of literature on video self-analysis and its positive effects on teacher development. They focused entirely on special education teachers, though at different stages in their career—pre-service, para-educators, and varying levels of in-service teachers. They wanted to know if video self-analysis was particularly helpful with certain populations or in certain contexts. The research team aimed to aggregate the growing body of literature on the effectiveness of video self-analysis specifically as a special educator development tool. They were also curious to see if the context of video analysis (e.g., self-contained or resource room) or participant characteristics (e.g., teaching experience, age) differentially impacted the effectiveness of video analysis. They conducted a meta-analysis of dissertations and peer-reviewed journal articles of single case special education video analysis studies with the primary research question: "What is the omnibus magnitude of effect of video analysis on the instructional practices of educators?" (p. 5).

Morin et al. (2019) included 33 single case studies of special educators using video self-analysis. They determined if those individual experiments met the What Works Clearinghouse (WWC) Single-Case Design Standards, such as a) including a manipulation of the independent variable (IV), b) reporting inter-observer agreement that meets a particular threshold, and c) a minimum of three attempts to demonstrate treatment effects for at least three points in time. With these rigorous standards, 18 of the studies met the criteria, though 17 of those were with some reservations. They proceeded to analyze the data for the entire group of 33 studies, as well as differentiating between the 18 that met the criteria and those that did not. They calculated the Tau-U effect size for each of the identified studies, as a total aggregate, and for each group of studies

according to whether they met WWC Design Quality Standards or not. Tau-U may be cautiously interpreted as having a small effect up to 0.62; a medium effect from 0.63 to 0.92, and a large effect from 0.93 to 1.00 (Morin et al., 2019).

The results of the meta-analysis studies can be seen in Table 3. Both the omnibus effect and the calculations for just the studies meeting WWC standards are reported for comparison.

Table 3

Study Groupings and Omnibus Effect Sizes (Morin et al., 2019)

Study Grouping	Number of studies	Tau-U (95% CI)	Tau-U Effect
All single case design experiments	33	0.85 [0.79, 0.91]	Medium
Studies meeting WWC standards	18	0.88 [0.81, 0.96]	Medium

The Tau-U effect sizes indicated a moderate positive effect size for video analysis for both the total group of studies and for the group that met the WWC Design Standards. The studies meeting the WWC standards had a slightly higher Tau-U with the 95% confidence interval Tau-U ranging from 0.81 to 0.96, which spans a medium to large effect size.

In addition to the omnibus effect, Morin and team categorized the video self-analysis studies into types to see if the context or participant characteristic made a

difference. They wanted to know if video analysis benefitted a particular educational setting, such as one-to-one versus small group versus large group instruction, or instructing communication skills, academic skills, or daily living skills. For participant characteristics, they created four subgroups including role (paraprofessional, preservice, and in-service), educational level (high school/GED, bachelor's graduate), age, and experience (none/first year, second or third year, and fourth year or more). Novice educators appeared to benefit the most from video self-analysis, which was a different finding from Fukkink et al.'s (2011) meta-study. Novice teacher case studies reached a large Tau-U effect size, 0.93 (95% CI [0.84, 1.00]), whereas more experienced educators had effect sizes in the moderate range, second/third year educators ES of 0.76 (95% CI [0.67, 0.86]), and four or more years of experience produced an ES of 0.85 (95% CI [0.78, 0.92]).

This meta-analysis was limited to special education teachers and primarily included studies that did not meet the WWC standards. However, the studies that met the WWC standards with some reservations and those that did not, yielded similar results, all with video analysis interventions having a positive effect. This study lends additional support for the usefulness of video self-analysis in developing educators and the importance of including an externally evaluated observation of the teacher. In particular, Morin et al.'s (2019) study suggested that the strongest results came for novice teachers, indicating that video self-analysis is an effective practice for teacher preparation programs. In contrast, Fukkink et al.'s (2011) meta-analysis did not produce evidence of a differential effect between novice and experienced practitioners, having both groups showing similar positive effects.

Both meta-analyses of video interventions provide support for using this as a tool to develop teachers. Video self-analysis provides a connection between the research that teachers learn about in their training and their actual K-12 classrooms. It allows them to reflect on their implementation of newly learned practices, and most importantly, the students' responses. Video analysis of their own classroom is genuine and authentic, not hypothetical or theoretical, and encourages candidates to closely examine student thinking. Video affords candidates the opportunity to reflect outside of the demanding minute-to-minute decision-making of teaching, and offers long-term follow up and comparison opportunities to see growth over time (Beisiegel et al., 2018; Derry et al. 2010; Knight, 2014).

Other Video Analysis Literature

In addition to the aforementioned studies, other research involves general video analysis, such as watching video of a mentor teacher or an example of model instruction (Beck et al., 2002; Blomberg et al., 2013; Blomberg et al., 2014; Seidel et al., 2013; Sherin & van Es, 2003; Sherin & van Es, 2009). Watching others instruct can also boost self-efficacy. Several studies emphasized the importance of peer video analysis teams where participants watch each other's videos and discuss the lessons (Beisiegel et al., 2018; Knight, 2014; Kimbrough et al., 2008; and Sherin and van Es, 2002, 2003, 2009). The focus of the discussion varies depending upon the video club's goals. Mentor teacher videos provide a similar teaching environment for candidates since they are working with the same students, school setting, and subject area. However, if candidates see themselves as very different from the mentor's personality, this may reduce the amount of perceived similarity. Beck et al. (2002) found that candidates who analyzed mentor teacher video

outperformed those who only observed the classroom live. They suggested that the video analysis group was able to focus more on student learning rather than superficial happenings, thereby reflecting more deeply and building mental models of how to perceive classroom activity. They argued that participants in the video group who had the advantage of rewinding, replaying, and putting on a new lens or focus question each time they watched a segment developed a more trained eye. Beck et al. (2002) postulated that this transferred to other instructional settings.

Sherin and van Es' (2003, 2009) research on practicing teachers' participation in video clubs yielded similar findings. In video clubs, groups of teachers watch each other's videos and use guiding questions to develop habits of focused reflection. Participants have reported that these habits transferred to their in-the-moment instruction. Examples for teacher prompts included: "What is the student saying about the learning? What did you see in the video about students' understanding of (fractions)?" (Sherin & van Es, 2003, 2009). One participant referred to this transfer of thinking as "video head" (Sherin & van Es, 2003). Teachers reported that the analytic mindset cultivated during the video clubs indeed influenced their classroom instruction (Sherin & van Es, 2003, 2009).

Several studies emphasized the importance of how EPPs structure the purpose and implementation of video analysis. The analysis activity should have guiding theoretical purpose (Derry et al., 2010; Seidel et al., 2013). Furthermore, EPPs should reduce the cognitive load for pre-service teachers during video observation so that they are not overwhelmed by the complexity of video nor caught in the trap of superficial description (Beck et al., 2002; Blomberg et al., 2013; Nagro et al., 2016; Sherin & van Es, 2003).

Short segments of video with guiding prompts renders viewing more manageable (Marzano, 2007). Blomberg et al. (2013) explained the novice tendency to describe events, whereas experts will more often reason, connect, and classify their observational insights. Video analysis, of self and comparable others, can help train the novice eye to become more sophisticated and expert. The literature review suggests that video analysis should be chunked into manageable parts with structured guidance, such as coupling the viewing with guiding questions or forms (e.g., rubrics or checklists) linked to effective practices. It also suggests that while self-analysis may be the most motivating for candidates (Beisiegel et al., 2018; Seidel et al., 2011), watching videos of similar others, such as trusted peers in an EPP, provides additional benefits, such as feedback and insight that the candidate did not notice (Beisiegel et al., 2018; Knight, 2014; Sherin & van Es, 2003, 2009). For these reasons, the current study counted both video analysis of self, and the quantity of video analysis teacher candidates provided others throughout their time in the EPP.

Chapter 3: Methodology

This study employed a non-experimental correlational design to determine if the quantity of video analysis conducted during the one-year EPP offers any predictive value for how a candidate will perform on the edTPA (Gall et al., 2007). The researcher collected data on 49 teacher candidates from two online cohorts, 2017-18 and 2018-19. The author planned to employ a multiple regression analysis to determine if the amount of video analysis conducted in the program accounts for unique variance in edTPA performance above and beyond writing skills. The literature review includes repeated evidence of a moderate, positive effect size for video analysis in teacher development. The literature tends to demonstrate a positive correlation, which could be due to publication bias preferring to publish statistically significant results. Since previous research suggests that strong writers have an advantage in scoring well on the edTPA, writing skills were included in the study (Goldhaber et al., 2017; Santagata & Sandholtz, 2019). The researcher conducted a correlational analysis using two years of data from 2017 to 2019, from the post-baccalaureate online teacher certification program at a university in Washington State. The author hypothesized a positive association between the predictor variables of WEST-B Writing scores and the number of videos analyzed within the yearlong program, and the outcome variable of edTPA scores.

Research Design

This correlational study planned to use a hierarchical regression to examine if the quantity of video analysis predicts unique variance in candidates' edTPA scores, after accounting for writing ability. Due to the literature review and previous results suggesting a positive relationship between video analysis and teacher performance, and

the presumed advantage of strong writers, this study aimed to determine if there is a significant relationship between these predictor variables and performance on the edTPA.

Research Questions and Hypothesis

1. What is the relationship between teacher candidates' writing abilities as measured by WEST-B Writing scores and teacher candidates' performance on the edTPA?
2. Does the amount of internship video analysis conducted during the program account for unique variance in predicting teachers' performance on the edTPA?

In light of the various studies indicating the benefits of structured video self-analysis, the author hypothesized a positive relationship between WEST-B Writing scores and edTPA performance scores, with additional unique variance accounted for depending upon the quantity of video analysis completed in the program. While Santagata and Sandholtz (2019) referenced a modest correlation between GPA and edTPA scores, WEST-B Writing scores offer a more targeted measure specific to the skill of writing. The edTPA involves a significant writing demand, and hence there is a common conception among EPP stakeholders that strong writing skills influence edTPA scores. Although the writing bias is a prevalent notion in the EPP community, the researcher did not find any empirical evidence to support it. In actuality, there is unpublished evidence that points to the contrary (cited in Whittaker et al., 2018). Since the author wanted to examine the relationship between writing ability and teacher performance assessment in a measured way and gather evidence for this commonly held conception, she chose to use WEST-B Writing scores rather than GPA. Hypothesizing

this common conception herself, the researcher intended to enter writing scores into the regression model first in a hierarchical fashion, assuming strong writers would have the most advantage. Next, the *number of videos critiqued* predictor variable would be entered into the regression to determine if these activities accounted for unique variance in the model (Tabachnick & Fidell, 2014).

Participants

Participant data came from graduate teacher education preservice students who enrolled in an online certification program at a private liberal arts university between 2017-2019. The university collects all of the data for program purposes. For the purposes of this study, these data are considered archival. A total of 49 students (n = 49) had edTPA scores and a data point for *number of videos critiqued*. The students represented various subject area endorsements, as shown in Table 4. Of the 49 candidates, 34 had WEST-B Writing scores due to the alternatives that were used to meet the basic skills program entrance requirements. For a regression analysis, a general guideline is to have at least 10-15 participants per predictor variable in order to detect an effect (Field, 2013; Gall et al., 2007). With two predictor variables, the lowest n = 34 meets this sample size recommendation.

Table 4

Intern edTPA Endorsements

Endorsement	Interns
Elementary Literacy	6
Elementary Mathematics	4
K-12 Performing Arts	1
Secondary English Language Arts	4
Secondary Social Studies	1
Secondary Math	3
Secondary Science	2
Special Education	19
Visual Arts	4
World Language	5
Total	49

Ten discipline endorsement areas were represented in this data set for edTPA scores. In Washington State, special education must be accompanied with a general education endorsement in order to earn certification. Thus, the special education candidates also had a general education endorsement, such as elementary, math, or English language arts. University program policy instructs interns pursuing dual endorsements to complete the edTPA in special education. To earn certification in Washington, interns need to pass the edTPA in one of their endorsement areas if they are seeking two subject area endorsements. Other students in the program may also have

chosen a dual endorsement pathway in the program, such as elementary and visual arts. Dual endorsing interns have internship placements in both subject areas, though there is usually a primary endorsement where the intern spends the majority of time. Interns typically complete the edTPA in their primary internship placement. Table 4 shows the endorsement for which the candidate completed the edTPA. In the literature review of video analysis, many of the studies focused on certain disciplines—special education (Morin et al., 2019; Nagro et al., 2016), math (Beisiegel et al., 2018; Santagata & Sandholz, 2019), science (Ceven McNally, 2015), and reading (Marker & D'Onfrio, 2015). The intent of this study was to focus on video analysis as a tool to develop teacher candidates generally, and not focus on a particular discipline.

The private university in this study operates a one-year accelerated program to teacher certification in both blended and entirely online formats. In the last three years, only the online programs have required the students to record their own classroom teaching for the purpose of supervisor coaching and evaluation. Thus, since the students in the online programs had access to video analysis and were required to record lessons for the program, it is this group of students whose data was used. Additionally, the online program cohorts had a similar programmatic experience compared to their blended, on-campus cohorts. The online students took coursework online from the same pool of online instructors and were assigned to a common set of online supervisors, thereby minimizing programmatic variance. However, teacher internship placements varied substantially, from urban to rural, public to private, and honors courses to self-contained special education classrooms. Moreover, each candidate had a different mentor teacher during internship. Thus, internship setting and assigned mentor are both potentially

confounding variables not included in this study, which does limit the findings. Other researchers have suggested that classroom setting and school socio-economic status impact edTPA scores (Greenblatt & O'Hara, 2015). Although the sites differed, the duration and expectations of the internship remained similar. Teacher candidates in these programs began a yearlong internship in late August and completed a full year's residency within the same classroom(s) through June.

The university assigns field supervisors to candidates for the duration of the year, making 10-12 observations and conferences throughout the year. Additionally, once a school field site identifies an appropriate mentor, the university confirms that the selection meets statutory requirements to provide each teacher candidate with regular on-site support and training. The field supervisors, with mentor teacher and candidate input, assess the teacher candidate's performance and growth throughout the year. Most of the interns in the online program live geographically far from campus, though within the state, thereby making face-to-face supervision a difficulty. Some supervisors live geographically closer to online teacher candidates, and are more willing to make a few of their 10-12 visits in person rather than a recording and teleconference. Some supervisors also have a stronger preference for live visits and a higher willingness and availability to travel. Thus, online candidates could conduct all 10 of their classroom observations via instructional recordings and teleconference debriefs, or do a blend of some traditional live visits with some instructional recordings. The blend of live and recorded observations varied between each candidate-supervisor arrangement.

Measures

WEST-B Writing

During the two years of data collection, candidates had to show both basic skills proficiency and subject matter proficiency by passing certain tests in order to earn Washington State certification. The basic skills tests included a battery of three tests assessing math, writing, and reading known as the Washington Educator Skills Test – Basic (WEST-B). Candidates could waive any and all of the three sub-categories with sufficient SAT or ACT scores, or state-approved equivalents from other states, which currently lists 18 possible options. Washington, as of 2019, no longer has a "cut score" for these basic tests, but still requires candidates to take them. During these two years of data collection, the passing score for WEST-B Writing was 240. Of the 49 identified candidates from the two years of cohorts, 34 had WEST-B scores. Other candidates had used prior college entrance exam scores or out-of-state teacher entrance exam scores as substitutions for the WEST-B Writing requirement. This study did not attempt to determine an equivalent scale for each of the possible alternatives for meeting the state's writing requirement for entering an EPP, nor did it attempt to find z scores for each test. The researcher did not have access to all of the alternatives. This study limited the correlation to WEST-B Writing scores. If a student took the WEST-B multiple times, the investigator input the first attempt in the analysis in an attempt to standardize the scores as much as possible for all candidates (i.e., using all first attempts).

Video Analysis Predictor Variable

Video analysis, the second predictor variable, included the number of instructional videos a candidate analyzed during the program (their own or a peer's video from the program). A video analysis counted as a data point if it included at least one analytical comment by the student regarding instructional matters in the video coaching

platform. For instance, video assignments that provided virtual tours of the classroom or were video journals explaining a candidate's particular practice, such as communicating with families, were not included. If candidates uploaded an instructional video for the supervisor but did not make their own comments, those videos were not included in the tally. The video coaching platform generates reports on how many videos a candidate submits for critique and how many videos they analyze. The researcher used the raw number of videos analyzed, and then subtracted the video submissions that did not include classroom instruction. For instance, if a candidate submitted a virtual classroom tour and made four comments on that video, the researcher subtracted that video from the participant's video analysis data. The number of videos analyzed was predominantly of a candidate's own classroom instructions as most of the activities were set to private, meaning only supervisors and candidates themselves had access to their videos. However, a few of the EPP assignment activities were set to peer review so that students could watch their similarly-endorsing peers' videos and give feedback to one another. These were also included in the video count as research suggests benefits from video peer review (Beisiegel et al., 2018; Kimbrough et al., 2008; Knight, 2014; Sherin & van Es, 2003, 2009; van Es & Sherin, 2002). All 49 candidates from the two cohorts who had earned a complete edTPA score also had a *number of videos critiqued* data point.

For the two years of program data included in the study, the requirement for student video self-analysis varied. Both cohorts were required to record their classroom instruction for observational purposes and share those recordings with their field supervisor. See Appendix A for the online supervision timeline that structured both cohorts. In the first year, 2017-18, candidates were encouraged to watch their own

recordings, but it was considered optional and primarily a source of observation for the assigned university supervisor. In 2018-19, online candidates were required to watch and analyze a minimum of two of their own videos and use evidence from those videos for two out of three written reflections. Students from both cohort years had to write three internship reflections, but only the second cohort had to integrate video analysis into those reflections. Both cohort years had students in certain courses who completed video peer analysis as assignments. Analysis for peer and self-review included timestamped comments and/or markers corresponding to the eight EPP standards. The university refers to their eight program standards, derived from Danielson's Instructional Framework (Danielson, 2013), as the Internship Performance Criteria (IPC). Video "markers" within the video coaching platform have been designed to align with these eight standards so that a viewer can timestamp a portion of the video to show evidence for evaluating a particular standard. The viewer can add comments explaining why a marker was chosen and how the evidence aligns with the expectation for meeting the standard or not. The markers mostly pertain to the first six of eight teaching standards. The final two standards usually pertain to competencies outside of immediate classroom instruction, *Family and Community* and *Professional Practice*. Supervisors gather other forms of evidence for these two standards, though there is an option to use these markers in case classroom video has relevance. The eight IPC standards are as follows, with their timestamp marker code used in the video platform in parentheses. For an illustration of what these markers and comments look like in the video platform, see Appendix C.

1. Expectations (Ex)
2. Instruction (In)

3. Differentiation (Di)
4. Content Knowledge (Co)
5. Learning Environment (Le)
6. Assessment (As)
7. Family and Community (Fa)
8. Professional Practice (Pr)

The eight IPC standards are broken into further sub-categories, each of which has a scale from 1 to 4 to evaluate the candidate's proficiency level: 1 = *Unsatisfactory*, 2 = *Basic*, 3 = *Proficient*, and 4 = *Distinguished*, similar to Fukkink et al.'s (2011) recommendation to use molar measures. The full IPC evaluation comprises 25 sub-categories and an 8-page document. Appendix D offers an example of one IPC category, *Expectations,* and its three sub-categories broken into molar rating scales. If a candidate sees an example of using a formative assessment during instruction, they might use the *As* (#6 Assessment) marker, and then write a free form comment explaining their thoughts such as, "6.2 Using white boards as formative assessment to see if students remember the vocabulary. Five students do not hold up any response. I move to the next question without any follow-up. Several other students have the wrong answer." The researcher wanted to determine if candidates who more often used this structured protocol for video self-analysis, and a similar analysis of peers' internship videos, would have a stronger performance of teacher competencies as measured by edTPA scores due to the influence of these video activities.

The researcher further disaggregated the video analysis variable into *number of self* and *number of other* videos that were critiqued in the video coaching platform. The

total quantity of instructional videos analyzed was broken down into how many times a student commented on their own instructional video (self), and how many peers' videos they commented on (other), combined as the total quantity of videos critiqued. Only a few activities in the video coaching platform were open for peer review for course assignments.

To gather the video analysis variable, the researcher ran reports from the video coaching platform to get the total number of videos critiqued—a function of the reports. Next, she subtracted video assignments that did not involve instruction. For instance, some of the video assignments in the coaching platform were titled *Introductions* and *Virtual Classroom Tour*. These video assignments had the dual purpose of familiarizing the students to the video coaching platform and technology, while also introducing themselves and their context to their field supervisors (*Virtual Tour)* and fellow cohort classmates (*Introductions).* The video coaching platform included video journal assignments that allowed student interns to share classroom artifacts on the camera while discussing their processes, such as their communication strategies with families. These video journals were of the interns alone. The researcher ran reports on each of these non-instructional video assignments. All of these non-instructional videos were subtracted from the *total videos* to create the variable of research interest, *total instructional videos analyzed*.

Once the *total instructional videos* count was established, the investigator calculated the number teacher candidates analyzed their own video (*self)* and the amount they analyzed others' videos (*other).* Students have varying program coursework, so not all students were required to conduct video peer review assignments. Most video

activities in the coaching platform were marked private for supervisory purposes. Thus, students did not have access to each other's videos unless it was explicitly for a peer review course assignment.

edTPA Scores as Outcome Variable

In Washington State, teacher candidates must pass the edTPA with a total score of 40 to earn certification (34 for languages). For most subject areas, there are five rubrics for each of three tasks—Task 1 *Planning,* Task 2 *Instruction* (the video component), and Task 3 *Assessment,* for a total of 15 rubrics. Table 5 describes the focus of each rubric for each of the tasks. At the time of data collection, Washington State was piloting three additional rubrics known as the *Student Voice* rubrics. However, since these scores were not part of the cut score for certification and were part of a pilot project, they are not included in the edTPA scores for this project. Each rubric has a highest possible score of 5, making 75 a perfect score. World languages (WL) and classical languages only have 13 total rubrics, and thus a lower possible score (65) and a lower passing score (34). Since the entirety of the language disciplines focuses on language itself, they do not include the specific rubrics that call out teaching components of academic language. With this passing score minimum, candidates need to earn an average of 2.67 on each rubric, though scorers only give whole scores, in order to earn certification. As with the WEST-B Writing skills, if a student took the edTPA multiple times, the first attempt with complete data was used as the data point. For instance, if a student received an error code for any of the rubrics, that score was not included in the data set; there was no mean substitution.

Table 5

Focus of Guiding Questions in edTPA Rubrics

Focus of rubric	Task
R1: How do plans build on each other	1—Planning
R2: Planning to support varied student learning needs	1—Planning
R3: Using knowledge of students to inform teaching and learning	1—Planning
R4: Identifying and supporting language demands[1]	1—Planning
R5: Planning assessments to monitor and assess learning	1—Planning
R6: Learning environment	2—Instruction
R7: Engaging students in learning	2—Instruction
R8: Deepening student learning/eliciting student responses	2—Instruction
R9: Subject-specific pedagogy	2—Instruction
R10: Analyzing teaching effectiveness	2—Instruction
R11: Analysis of student learning	3—Assessment
R12: Providing feedback to guide learning	3—Assessment
R13: Student understanding and use of feedback	3—Assessment
R14: Analyzing students' language use[1]	3—Assessment
R15: Using assessment to inform instruction	3—Assessment

[1]These are the two rubrics which are absent from the world and classical languages.

Pearson, an external organization, scores the edTPA portfolio. Approximately 10% of portfolios are double-scored, and Pearson reports multiple reliability coefficients that indicate a high level of internal scoring consistency, such as Cronbach's alpha of .91 (Pecheone et al., 2016). Four of the Task 2: *Instruction* rubrics use the artifact of

classroom video evidence, while the fifth rubric in Task 2 focuses solely on the candidate's written reflection and analysis of how the lessons went. The total score on 15 rubrics determines passing or failing in Washington State. However, since those in world and classical languages have two fewer rubrics, the data analysis required an imputation technique to estimate those absent values. The researcher chose personal rubric mean substitution as the preferred method. The average rubric score for each world language (WL) student's portfolio was added twice again to their total score to make it equivalent to a 15-rubric, 75-point possible score. Deleting the WL students (n = 5) from the data set was not a desirable choice since the researcher wanted to include them as fully participating students in the online cohorts. The WL edTPA simply leaves out the two rubrics on academic language since the entire subject involves teaching communicative language skills; they are not a random subsample. Thus, the researcher chose mean substitution for this non-existent data using the students' own mean rubric scores, not group means. For example, if a world language participant earned a 39 over 13 rubrics, that averages to a 3 for each of the 13 rubrics for that particular student. A dummy score of 3 was placed into those two missing rubrics for academic language and added to the total score (39 + 6 = 45). Mean substitution as a method for estimating missing values brought the total possible maximum score to 75 over 15 rubrics. Each of the five WL students had their own personalized mean substitution added twice to their scores in order to retain what variability existed among students. While the edTPA is reported in whole numbers, the researcher rounded to the nearest tenth for this imputation technique (e.g., a WL student's total score of 42 divided by 13 rubrics would result in a 3.2 mean being added to the two missing rubrics). Personalizing the mean substitution as a form of prior

(current) knowledge about the students' own performance and rounding to the nearest tenth were both done to preserve some of the variability within rubric scores (Tabachnick & Fidell, 2014). Mean substitution is considered a conservative estimation technique for missing values because it reduces the possible variability (Tabachnick & Fidell, 2014).

In addition to the composite edTPA score, the researcher disaggregated the edTPA variable into the three tasks: Task 1 (Planning), Task 2 (Instruction), and Task 3 (Assessment). Each of the tasks has five rubrics, with a maximum score of five on each, for a possible total score of 25 on each task. The researcher planned correlations between the disaggregated variables of *self* and *other* instructional videos critiqued, and the three separate edTPA tasks, as well as correlations with the composite scores of the total number of instructional videos critiqued during the program and the total edTPA composite score. By disaggregating the variables, more information could be gathered regarding the relationship between a particular type of video analysis—*self* or *other*—and its relationship to the particular edTPA task (planning for learning, instructing, and assessing student learning).

Chapter 4: Results

The researcher collected the relevant data on the 49 teacher candidates from two online cohorts, 2017-18 and 2018-19. Descriptive statistics for all of the aggregate variables are shown in Table 6. The literature review suggested a small to medium positive effect on externally evaluated (compared to self-evaluated) teacher competence when using structured video analysis in the development of teachers (Fukkink et at. 2011; Morin et al., 2019; Nagro et al., 2016; Santagata & Sandholtz, 2019). The researcher hypothesized that she would find a small to medium positive correlation with edTPA scores due to the existing research.

Table 6

Descriptive Statistics of Variables

	n	Range	Min.	Max.	Mean	*SD*	Skew	Kurt
WEST-B	34	68	229	297	268.88	17.44	-.34	.78
edTPA	48	24	32	56	45.99	5.76	-.51	-.14
Videos	48	17	0	17	6.5	4.90	.45	-.63
Critiqued[1]	(49)	(29)	(0)	(29)	(6.96)	(5.82)	(1.27)	(2.88)

[1]The row for *videos critiqued* includes the descriptive statistics with and without the identified outlier.

Preliminary Data Analysis

Before conducting the regression analyses, the researcher calculated descriptive statistics to analyze the appropriateness of the data. The variable, *number of videos critiqued*, flagged concern due to the high statistics of skewness (1.27) and kurtosis (2.88) outside of the recommended range of ±1; the closer to zero, the better (Field, 2013).

Since the sample size is relatively small, the researcher calculated *z*-scores by dividing the skewness by the standard error of the skewness and followed a similar procedure for kurtosis. These calculations produced *z*-scores for skewness (3.74) and kurtosis (4.31). A perfectly normal distribution would have zero as a *z*-score. Field (2013) explains that an absolute value above 1.96 for *z*-scores in these two areas indicates deviation from normality that is statistically significant at the $p < .05$ level, and an absolute value greater than 3.29 is significant at the $p < .001$. Both skewness and kurtosis for *number of videos critiqued* fell above the 3.29 *z*-score threshold with all data points included. Figure 1 shows a scatterplot of edTPA scores and *number of videos critiqued* with a visible outlier with 29 *videos critiqued*.

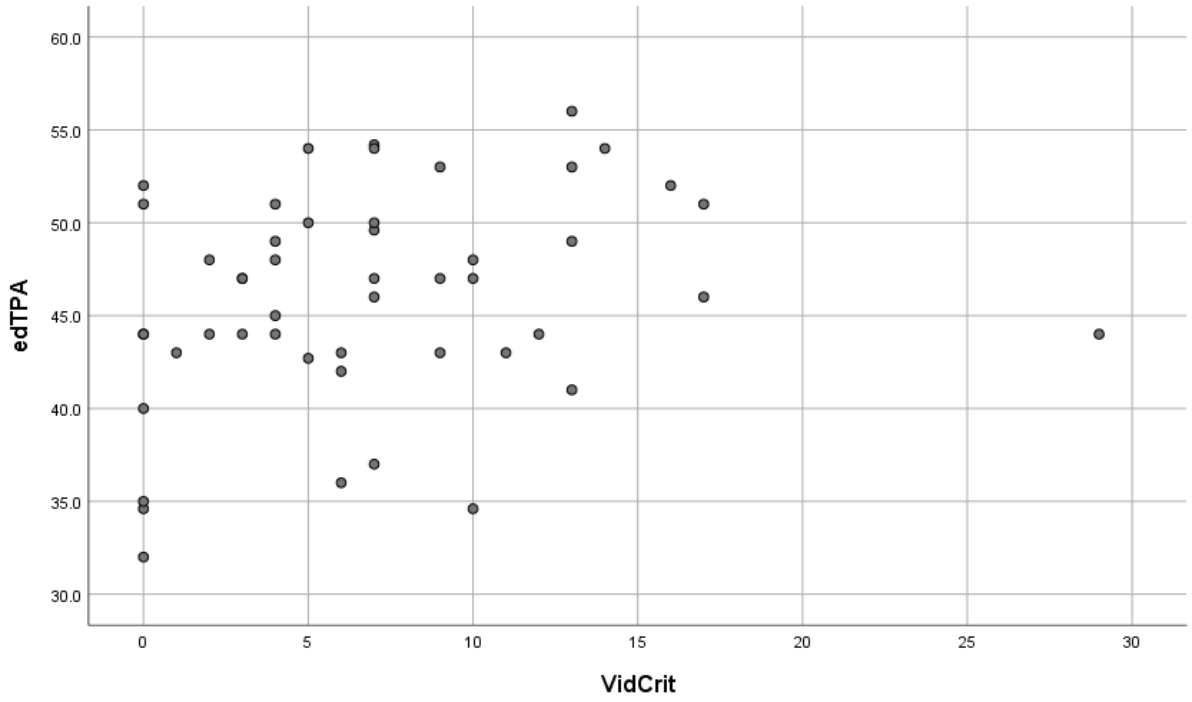


Figure 1. Scatter Plot of "Number of Videos Critiqued" with Outlier

There are multiple ways to determine outliers. For this study, the researcher employed the scatterplot and standardized residuals. In order to identify statistical outliers, she looked for any data point that fell outside of the recommended 3.26 standardized residual range (Field, 2013). The researcher ran standardized residuals for edTPA scores, WEST-B Writing, and *number of videos critiqued*. The data point of 29 *videos critiqued* had a standardized residual of 3.79, showing that it indeed fell outside the recommended three standard deviations from the mean, or greater than 3.26 standardized residuals (Field, 2013). The other variables, WEST-B, and edTPA scores, had data points within the acceptable range of standardized residuals between -3.26 and +3.26. The outlier was removed from the data set to see if this improved normality. When the outlier was taken out of the analysis for *number of videos critiqued*, the *z*-scores for skewness and kurtosis both fell into the recommended range below an absolute value of 1.96; skewness *z*-score without outlier .45/.34 = 1.32 and kurtosis *z*-score -.63/.67 = -0.94. The researcher excluded the one identified outlier variable from the analyses by defining it as "missing data" in SPSS. Doing this improved the skewness and kurtosis of the *videos critiqued* variable as seen in Table 6, which includes the descriptive statistics with and without the outlier.

To examine the associations between variables, the investigator conducted correlation analyses on each of the predictor variables with edTPA scores. Since each variable had a differing quantity of data points, the researcher ran the correlations separately in order to use all of the data points available for that particular variable. To illustrate, of the 48 students with *number of videos critiqued*, only 34 also had a WEST-B Writing score. At this preliminary stage of bivariate correlation analysis, it was clear that

the hypothesized positive relationship between writing scores and edTPA performance was unsupported. There was no evidence in this data set showing a relationship between WEST-B Writing and edTPA scores, $r = -.004$, $p = .984$. However, there was, as hypothesized, a statistically significant positive association between *number of videos critiqued* and edTPA scores.

Table 7

Correlations of Predictor Variables to edTPA Performance

Predictor Variable	*n*	Pearson's *r*	*p*-value
WEST-B Writing	34	-.004	.984
Videos critiqued	48	.342	.017

Relationship between Writing Ability and edTPA Performance

The first research question focused on the relationship between teacher candidates' writing abilities as measured by WEST-B Writing scores and teacher candidates' performance on the edTPA. There was no statistically significant relationship between those two variables as the correlation was near zero. The researcher failed to reject the null hypothesis. In this data set, academic writing ability was not a significant predictor of edTPA performance. Thus, the researcher determined that continuing with the original plan of a hierarchical regression would be illogical.

Students who met the basic writing entrance requirement another way, such as with strong high school SAT or ACT scores, were not included in the correlational analysis. To determine if there was a difference in edTPA performance for students who used the WEST-B Writing for EPP entrance and those who did not, the researcher

performed an independent samples t-test. The results of the Levene's test, $F(1, 47) = .008$, $p = .931$, indicated that the two groups had equivalent variance. On average, students who used an alternative to the WEST-B Writing requirement, $n = 15$, ($M = 48$, $SE = 1.39$) did perform better on the edTPA compared to those who used the WEST-B, $n = 34$, ($M = 45$, $SE = 0.99$). However, this difference of 3 points, 95% CI [-6.44, 0.62], was not significant $t(47) = -1.66$, $p = .10$. It represented a medium-sized effect, $d = 0.54$.

Relationship between Video Analysis and edTPA Scores

The second research question focused on the quantity of internship video analysis conducted during the program and its predictive value on teachers' performance on the edTPA. Since the WEST-B Writing scores in this study were not predictive of edTPA performance, a hierarchical regression was not deemed necessary. A linear regression with *number of videos critiqued* as the predictor variable and edTPA as the criterion variable showed that the amount of video analysis conducted in the EPP accounted for 11.7% of the variance in edTPA scores, $R^2 = .117$, $p = .017$.

The researcher conducted correlational analyses on the disaggregated variables of video analysis of self and others as well as the three separate edTPA tasks. The resulting correlation coefficients are shown in Table 8. Several correlations reached statistical significance, including the one between the total number of videos critiqued and the total edTPA score. However, an even stronger correlation, significant at the $p < .01$ level, emerged for the number of *other* (peer) videos a student critiqued and the total edTPA score. Task 3, *Assessment,* showed the strongest correlations of the three edTPA tasks with the frequency of video analysis in all three video analysis categories: *total number, number of self-critiqued,* and *number of other critiqued*. The strongest positive

correlation, significant at $p < .01$, was between the total number of other videos students analyzed and their own total edTPA score ($r = .367$, $p = .009$). The number of videos students self-critiqued did not reach the threshold of statistically significant correlations with any of the separate edTPA tasks, although there is some evidence of a positive relationship with the composite edTPA scores ($r = .268$, $p = .066$).

Table 8

Correlation Table of Disaggregated and Composite Variables

	Task 1	Task 2	Task 3	edTPA Total
Number of self-critiqued	.209	.180	.239	.268[1]
Number of other critiqued	.266[1]	.267[1]	.335*	.367**
Total videos critiqued	.248	.181	.355*	.342*

Note. ** $p < .01$ *$p < .05$ [1] $p < .07$

Secondary Findings

The researcher's hypothesis and the common conception that writing ability predicted edTPA performance was unsupported. Writing ability as measured by the WEST-B Writing scores had no significant relationship to edTPA scores, $r = -.004$, $p = .984$. The results from the correlational analyses showed that the number of videos analyzed had a positive relationship with edTPA scores. This sparked an investigation into the difference in performance of these two cohort years since one cohort year was required to do more video analysis than the other. The researcher explored if there was a difference in performance between the two cohort years, 2017-18 and 2018-19, since each cohort had naturally occurring programmatic expectations that differed in terms of video self-analysis requirements.

Difference between Cohort Years and Alternate Expectations

Since the data included two online cohorts, 2017-2018 and 2018-2019, with differing programmatic instructions, an exploratory independent samples *t*-test was conducted to see if there was a difference between the two cohort groups in edTPA performance. The cohorts experienced slightly different programmatic requirements regarding video self-analysis. The first cohort who had access to the video analysis software was told that self-analysis was encouraged, but optional. The second cohort year, 2018-2019, was directed to analyze their own instructional videos at least two times and to use those videos as evidence in their teacher reflections. In both cohort years, teacher candidates were required to write three internship reflections about their classroom practice, but the EPP did not prescribe the source of evidence for those reflections for the 2017-18 cohort (i.e., classroom artifacts might include lesson plans, student work samples, emails to parents, or recorded lessons). Table 9 shows the descriptive statistics for both online cohort groups. Although the mean differences were not statistically significant, the 2018-19 cohort did conduct more video analysis overall, both for self and other.

Table 9

Two Cohort Years' Means

	n	*M* (edTPA)	*M* (Number Self-critiqued)	*M* (Number Other-critiqued)	Total Critiqued
2017-2018 Cohort	27	44.6	3.00	1.89	5.44
2018-2019 Cohort	21	47.7	4.52	3.36	7.86

On average, the 2018-19 cohort which was required to conduct self-analysis at least twice ($M = 47.7$, $SD = 4.56$), performed better on the edTPA than those from 2017-18 who were not required to do any self-analysis ($M = 44.6$, $SD = 6.33$). Although the difference was not statistically significant, there was some evidence of an effect with the compulsory video self-analysis cohort, $t(47) = -1.927$, $p = .060$. This difference represents a medium-sized effect, $d = 0.56$ (Field, 2013).

Chapter 5: Discussion

Substantial literature exists that video analysis of authentic classroom teaching merits a place within EPPs. This correlational study adds empirical support to this claim. Teachers find video analysis of classroom teaching motivating and authentic (Beck et al., 2002; Beisiegel et al., 2018; Schieble et al., 2015; Seidel et al., 2011). Moreover, teacher competence as measured by external evaluations also appears to improve with structured video analysis. Evidence shows that the manner in which the video analysis is structured influences the size of the effect (Fukkink et al., 2011; Nagro et al., 2016). Since the present study did not include a formal intervention as an experiment, nor required protocols for the video analysis, it is possible that a medium to large effect could emerge if the university designed a carefully structured video analysis protocol, as Morin et al. (2019) found for novice teachers.

Bias towards Strong Academic Writers in edTPA Unsupported

Although there is a common conception of strong academic writers being favored in the edTPA, the researcher did not discover an empirical literature base to support this. It seems an intuitive relationship since the edTPA requires a lot of writing and commentary, but the researcher did not find empirical support for this idea in the literature review nor in the correlational analysis. A few studies referred to this relationship (Greenblatt & O'Hara, 2015; Kim, 2019; Santagata & Sandholtz, 2019), but none actually provided empirical evidence. One report identified this particular claim as a critique of the edTPA (Whittaker et al., 2018), and they refuted this argument by pointing out that there is actually no empirical evidence to support this idea. Whittaker and team (2018) referred to an unpublished study that showed similar correlational results as the

present study, meaning there was no predictive value nor relationship between standardized writing scores and edTPA performance. This correlational study lends support to deflate that argument. The correlation was essentially zero, indicating a lack of relationship between the two. The edTPA developers train scorers regarding this potential writing bias and inform scorers to be aware of this potential partiality. Scorers are instructed to mark portfolios according to the rubric language and not take off points for grammatical errors or weak writing. The assessment also permits bullet point explanations in the commentary sections (Whittaker et al., 2018). This study provides data that challenges the current conception held in many EPPs that the edTPA is biased towards strong writers. Perhaps academic writing skills are not the conflating variable that EPPs imagine.

While the correlational analysis did not provide evidence of a positive relationship between strong academic writing and edTPA performance, the *t*-test of the two groups who did and did not use the WEST-B Writing as an entry requirement indicated some difference. The medium positive effect size in favor of those who had used alternative writing proficiency scores to enter into the EPP, $d = 0.54$, raises the possibility of something going on. While the difference between the two groups was not significant, the effect size is substantial enough to cause curiosity. It is possible that the population of students who took the WEST-B Writing differed academically from those who used alternative scores to waive the writing requirement. For instance, students who attended four-year universities as freshmen typically had SAT and ACT scores, whereas many community college transfer students did not. Those who used their SAT or ACT scores to waive the writing requirement were those who had sufficiently high scores.

Santagata and Sandholtz (2019) did cite undergraduate GPA as having a small, positive correlation with edTPA performance. It is possible that not including "waiver" students in the correlation removed many strong writers from the analysis and limited the correlational finding. If the populations did differ, it may be that a group of students who had stronger academic writing backgrounds was not included in the sample because they were essentially "excused" from the WEST-B Writing requirement by meeting it through alternative ways. However, there were several ways to waive or substitute the WEST-B Writing requirement, which is why those alternatives were not included in this study. The investigator did not attempt to find WEST-B score equivalence with each option in order to include those students' data in the correlational analysis. There are several possible writing proficiency substitutes including out-of-state EPP entrance exams, previous PRAXIS scores – a former Washington State requirement, SAT writing scores, or ACT scores. Some students took the SAT when it did not include a required writing portion, so the year a student applied to undergraduate university influenced the possibility of having a substitute score.

Of the 49 students, there were 34 students accounted for in the WEST-B Writing variable. Scores ranged from 229 to 297 upon the first attempt, with 300 being the maximum score possible. This shows a decent amount of variability in academic writing ability among the included group as measured by this standardized test. The data captured from these cohorts for a correlation analysis did not provide evidence of academic writing as a predictor of edTPA performance. This contradicts a very common conception held within EPPs, but does not contradict other empirical findings from the literature review (Greenblatt & O'Hara, 2015; Kim, 2019; Santagata & Sandholtz, 2019).

The literature review produced a noticeable lack of empirical support for this conception. The lack of support for writing predicting edTPA performance may provide some confidence that the edTPA does not conflate writing ability with teaching competence. Rather, this empirically unsupported relationship adds to the reliability of the assessment as a measure of teacher competencies as Whittaker and team (2018) assert in their rebuttal to edTPA critiques.

Critiquing Self Versus Similar Others

Literature on video analysis in the development of educators includes both analyzing self and others. This particular study focused on video analysis of either self or similar peers, such as teacher video clubs or professional development groups (Beisiegel et al., 2018; Sherin & van Es, 2003, 2009), mentor teachers (Beck et al., 2002), or similar age and subject peers (Seidel et al., 2011). In other words, the focus of this study was not on a bank of instructional videos used as demonstrations or exemplar best practices. The videos of interest included everyday classroom interactions of the teacher candidates themselves or those very similar to their context, other teachers whom they often knew from the program. Focusing on these two areas—self and similar others—highlights Bandura's two areas for building self-efficacy. He asserts that one's own mastery successes build self-efficacy the most. Additionally, Bandura claims that people can boost their own self-efficacy vicariously by watching similar others succeed. Accordingly, it may be that as teacher candidates watch themselves in their own classroom and notice particular successes, they are more likely to repeat those actions. For example, they can notice how setting up a particular communicative activity in world language with student demonstrations, word banks, and sentence stems gets more

students actively participating in the target language with one another. Similarly, they can watch a peer student teacher conduct similar activities and believe in their own abilities to implement those same strategies in their classroom, an example of Bandura's vicarious success.

In the literature review, self-analysis had modest advantages over viewing others when the two were compared (Beisiegel et al., 2018; Seidel et al., 2011). However, it was also noted that analyzing and discussing videos with peers had benefits (Beck et al., 2002; Beisiegel et al., 2018; Knight, 2014; Sherin & van Es 2003, 2009; van Es & Sherin, 2002), in particular the ability to critique or notice salient aspects of instruction the teacher missed. For these reasons, the investigator chose to include the frequency of both kinds of video critique in her correlational study and to disaggregate by video critique type—*self* or *other*. There was an overall positive relationship between the frequency of video analysis conducted during the EPP and the composite edTPA score. When the correlations were disaggregated by type, the frequency of critiquing *others*' videos emerged as a stronger predictor than the quantity of *self*-critiques a candidate completed. In fact, the strongest correlation came from the number of *other* videos a student had critiqued and their total edTPA score ($r = .367$, $p = .009$). This results in an effect size of $R^2 = .135$. The second strongest correlation came from the quantity of *total videos critiqued* and Task 3: *Assessment* ($r = .355$, $p = .013$), $R^2 = .126$. This was somewhat surprising as the researcher predicted Task 2: *Instruction* to have stronger correlations since that task includes video analysis of classroom instruction. The edTPA Task 3: *Assessment* has candidates analyze the K-12 performance data for the whole class on the lesson segment's culminating assessment. In addition to whole class achievement data

and patterns, teacher candidates select work samples from three focus students to represent varying student progress. The submitted work samples in Task 3 include the students' work along with the teacher candidate's feedback after grading. It could be that the attention to feedback and using a rubric to analyze a peer's video attuned the teacher candidates to look for specific evidence in artifacts. Honing their analysis and feedback skills with their peers via video analysis may have positively influenced their ability to analyze their K-12 students' learning and give constructive and supportive feedback. In the peer review video analysis assignments for coursework, there were criteria included, and students were instructed to align their feedback to those criteria. Not only did teacher candidates practice spotting evidence and aligning their feedback to criteria in peer analysis, but by virtue of participating in these peer reviews, they were also receiving feedback from peers on their own videos. Thus, they had the advantage of others' perspectives of what went on in the instructional moment, but not from a high stakes evaluator, which can be more threatening or off-putting (Beisiegel et al., 2018; Knight, 2014).

Albeit unexpected, the strong correlation between *critiquing others* and Task 3: *Assessment,* offers hopeful possibilities. Pecheone et al. (2016) have noted that candidates across the nation have tended to score lowest on the *Assessment* task. However, Newton (2010) in his value-added measure study of PACT, edTPA's precursor, reported the stronger predictors for future K-12 student achievement came from candidates scoring high in assessment and describing student language development (part of academic language rubrics in the edTPA). Teacher educators desire practices that produce more proficient assessors of student learning. This correlational analysis points to a potential

tool that EPPs can use to foster candidates' growth in the area of student assessment: video peer review.

Simulating a Form of Video Clubs—Critiquing Self and Other in Same Activity

The secondary finding between the two naturally occurring cohort groups, 2017-18 and 2018-19, in the EPP provided additional evidence that using video self-analysis to help developing teachers hone their craft has empirical support. There was an effect size of a Cohen's $d = 0.56$ in favor of the cohort that was instructed to review and reflect on at least two of their own instructional videos, which is considered a medium effect size. Ideally, the teacher candidates were expected to interact with their field supervisor, mentor, or a peer review group regarding their video analysis, and many did. However, the interactions between students and others' comments on the same instructional video were not a focus of data collection in the current study, and cannot be expanded upon. The *critiqued other* disaggregated variable showed stronger correlations than the *self-critiqued* frequency. However, the only way students could have analyzed one another's videos was if the instructional videos had been uploaded to a "peer review" permitted activity. For these coursework assignments, students were asked to comment on their own and like-endorsement peers' videos (e.g., students pursuing an endorsement in world language would comment on each other's videos for a class learning activity). If candidates did not share the same endorsement, they were grouped with peer interns teaching similar endorsement and grade level, such as high school social studies with high school English language arts since both are humanities disciplines with a strong focus on reading and writing.

Sherin and van Es (2003, 2009; van Es & Sherin, 2002) have conducted research on video clubs as professional development for practicing teachers, which is similar to these video analysis activities in the EPP. Members of video clubs use the classroom artifact of video evidence to discuss specific topics and learn from one another's practice. Sherin and van Es described a *professional vision* that is developed when teachers look at what actually happened – how students and teachers really responded to classroom events. Teachers, or their colleagues, might notice a missed opportunity to ask a probing question to help a student learn more deeply or challenge a misconception, or they might notice an effective move that enhanced student learning or engagement. The video club articles by Sherin and van Es were not prioritized in this literature review because they involved practicing teachers and did not have external, empirical measures of teacher competence. However, the results of this correlational study combined with the Beisiegel et al. (2018) study that compared different kinds of video professional development groups, show that these kinds of *self* and *other* (colleague) peer review or video club development groups are a promising and well-received practice.

In video clubs, teachers can identify a learning principle together, and then look for that particular principle in their own recordings, such as having a 5:1 ratio of positive to corrective interaction with students, or categorizing types of questions asked (Derry et al., 2010; Knight, 2014). As another example, world language teachers might identify a language acquisition strategy such as a scaffolded listening activity—what to do the first time through listening, the second time through, and so on. The videos allow the teachers to see their own implementation of the strategy, compare to one another, and point out areas of lesson success or growth opportunities. This type of video analysis is actually a

form of both self-critique and other-critique; they are entwined. Since the goal of educator preparation is to teach ways of thinking and habits of mind for teachers to take into their professional careers, this video club simulation during the EPP offers an effective tool that interns can bring to their own future classrooms and learning communities. Bandura's (1982) social cognitive theory asserts that self-efficacy and motivation to perform certain actions depend upon the reciprocal factors of the environment, one's own internal personal factors, and behaviors. It was stated earlier that EPPs try to influence the teacher behaviors since that is where they have the most opportunity for influence. Training teachers in looking for specific timestamped evidence of classroom activity is one of these behaviors that can influence teacher habits and ways of thinking into the future. Noticing classroom application of a particular teaching standard or principle, or lack thereof, is analogous to training young K-12 writers to support their claims with evidence from the text. Teacher educators train teacher candidates to ask their students: *Where do you see evidence of this? Support your claim.* Teacher educators must do the same with the teacher candidates themselves and not take their reflective word for it.

Limitations

This is a correlational study and thus, cannot establish causation (Field, 2013; Gall et al., 2007). Further research is needed using an experimental design with video self-analysis or video clubs of self and peer analysis in EPPs. There are other confounding variables thought to influence edTPA performance scores which are not included in this present study, such as teacher technical savvy, the socioeconomic status of the school, and the skill of the mentor teacher with whom the teacher candidate is

working (Greenblatt & O'Hara, 2015; Santagata & Sandholtz, 2019). Additionally, the teachers throughout the two cohort years did not have identical program experiences. There are several course sequences depending upon endorsement choice, such as a track for elementary endorsers, a track for math endorsers, and yet another for special education. The university edTPA support team modifies the resources that they offer year to year due to program evaluation and adjustments. Many of the activities that were intended to assist candidates in putting together their edTPA portfolios were optional. Several support activities were embedded in the program's coursework, but there were also optional workshops provided by the university. Students took advantage of these optional supports to varying degrees, which could be another contributing factor to edTPA performance. The sample size in this study was also relatively small, which could contribute to a Type II error.

In addition to the confounding variables, there was also a fair amount of imprecision in the video analysis and an absence of measuring K-12 student learning. Only the frequency of different videos analyzed counted in this correlational study. Thus, if students commented one time on their own video or 30 times on that same video, it simply counted as one video critiqued. It would be helpful for future research to delineate the type and nature of video feedback that student interns find most valuable and which have the most impact on teacher competence and skill. Teacher perceptions are important for developing effective and well-received activities, but it is also important to measure classroom implementation and student learning. This study does not include K-12 student learning outcomes and cannot provide information to that end.

The type of video analysis quantified in the current study varied. While students were told to use the Internship Performance Criteria evaluation tool used by the EPP for student teacher instruction or choose other discipline-specific foci, there was not an attempt to verify the quality or fidelity of student comments in the correlational analysis, nor quantify back and forth interactions among multiple viewers. Rather, the investigator collected data only on whose video was the object of analysis: *self* or *other.* If students commented on a video of their own, it counted as *self,* and if they commented on someone else's video, it counted as *other*. Thus, the comments could vary substantially in quality, depth, length, and interactional engagement from peers or a supervisor. Appendix E shows a few illustrative comments from teacher interns. Some comment strings had several interactions between the teacher intern who uploaded the video and either the field supervisor or a peer reviewer, while other videos may have had just one commenter and no documented interactional analyses in the video coaching platform. Some comments may have focused more on the K-12 student actions and words in the video, while other comments focused more on the teacher behaviors, and some a blend of both.

While there was a skeletal structure to the video analysis using the eight Internship Performance Criteria described earlier, the correlational analysis only incorporated the frequency. The actual quality and nature of video analysis comments were not accounted for. There was likely great variability in the depth, length, and type of comments students made. For instance, some teacher candidates may have chosen to focus on student learning and their responses to the teacher and one another. Another candidate may have analyzed the teacher's behaviors. This correlational study does not offer any insight as to the kind of critique that students conducted, nor which kind had

more predictive value on teacher skill as measured by the edTPA. In Marker and D'Onfrio's (2010) study of reading teachers, they found that the real value of video came with the dialog between reviewers after watching and sharing strategies in a collegial format. The teachers' initial adverse reactions to videoing themselves persisted with personal written reflections, but transformed into a positive experience once the participants perceived the process as more collaborative and collegial, rather than evaluative.

Implications

Several studies have important implications for EPPs, primarily that guided video analysis has positive effects. In Nagro et al.'s study (2016), all student teacher candidates participated in video self-analysis, and both the comparison and treatment groups improved on measures of self-perceptions of teaching ability and actual instructional skills as measured by an external scorer, though internal to the program. However, the more structured video self-analysis group outperformed the less structured group. If developing teacher self-efficacy and habits of mind that transfer to classroom instruction is a goal of EPPs, then planning intentional video analysis during student teaching may help facilitate this end. The caution with self-reporting is that perceived ability and actual ability do not always align as Nagro's team (2016) demonstrated with the mismatch of self-reported ability and outsider scored observations. Self-reporting biases can threaten the accuracy of data (Field, 2013; Gall et al., 2007). Thus, having an outsider scoring actual teaching remains an important component, rather than relying solely on student self-reporting. Non-institutional, outside scoring minimizes further bias and is one reason this study used external edTPA scores. The structure of the video analysis focus and

feedback is important for both the intern and the feedback provider. External encouragement by a peer or supervisor could be considered a Bandura "efficacy builder" activity and help maximize video feedback impact.

Using structured forms and guided prompts have been shown to increase the impact of video analysis (Derry et al., 2010; Fukkink et al., 2011; Marker & D'Onfrio, 2010; Nagro et al., 2016; Sherin & van Es, 2009). Cognitive load theory asserts that narrowing down explicit expectations for a task will improve performance on that task. Chunking video analysis into manageable parts and guided steps mirrors the strategy taught to teacher candidates to use in their own P-12 lessons (Blomberg et al., 2013; Marzano, 2007; Seidel et al., 2011). Chunking limits the over-taxation on working memory where humans process information (Marzano, 2007; Medina, 2008). Other studies have shown that focusing viewer attention, often multiple times with varying lenses, can deepen the reflection and attune the teacher to student learning rather than superficial descriptions (Beck et al., 2002; Derry et al., 2010; Schieble et al., 2015; Sherin & van Es, 2003). Derry et al. (2010) compared this unprecedented close documentation capability of video analysis to a powerful microscope. They recommend that each viewing begins with clear and theoretically motivated questions so that the viewer retains perspective and refrains from getting lost in the myriad complexities of classroom interactions. For instance, the following guiding questions illustrate theoretically motivated questions:

- *How do you think your words positioned your students as readers and writers?* This study integrated positioning theory with a focus on discourse

analysis and interactional awareness within the classroom. (Schieble et al., 2015)

- *How does your lesson connect students' prior knowledge to the new concepts? How do you anticipate misconceptions?* These questions focus on modern learning theory and Piaget's theory of building upon known schema. (Danielson, 2013; NASEM, 2018)

Principles of andragogy tell us that adults are self-directed and goal-oriented and desire hands-on training that have immediate application value (Merriam, 2001; Zepeda, 2012). Video analysis in this context allows interns to choose their focus, select their goal, and then apply it to their future classroom lessons with the opportunity to follow up.

While this study cannot claim causality, there is some evidence of a positive relationship between video analysis and teachers' ability to assess students' work. Santagata and Sandholtz (2019) reported a statistically significant positive correlation, $r = .237$, $p < .05$, between candidates' mathematical content analysis and the assessment component of the PACT. In the present study, there was a positive correlation between the quantity of video analysis and Task 3 of the edTPA, *Assessment.* In both of these correlational studies, it was the assessment component of the performance portfolio that had the strongest positive relationship with video analysis. It may be that practice with aligning feedback to a particular rubric has given teacher candidates additional practice at targeted assessment. For teachers who may lack skill in giving specific and focused feedback, video analysis with a structured form, such as the IPC in the present study or the *Classroom Video Analysis—Mathematical Content* in Santagata & Sandholtz (2019), may help develop teachers' assessment lens. While the rubrics differ depending upon

content area and age level of the K-12 classroom, the procedural skill of attending to rubric language and highlighting specific evidence to support a particular score during video analysis may be transferred to the K-12 context. Teachers learn to narrow their focus and align feedback. Hattie and Timperley (2007) discuss the importance of specific and targeted feedback to close the gap between current and desired performance or understanding. However, some teachers may struggle with providing specific feedback and fall back on generalities such as “Good job” or “Needs work.” Video analysis that is accompanied by guided rubrics or checklists offers a practical way for teachers to practice identifying a learning outcome and then providing specific feedback to help lessen the performance gap.

Strengthening this assessment lens also gives teachers the opportunity to turn that lens towards themselves. Many states have adopted teacher evaluation systems that include various pedagogical rubrics. Teachers may choose a particular evaluation tool of relevance to their context, such as the Danielson Framework (Danielson, 2013), select a personal growth area, and narrow their focus to that particular evaluation item and collect video evidence for that rubric to demonstrate growth over time. Developing an assessment lens necessitates identifying a particular outcome and way to measure success. Video analysis with structured forms provides a process for narrowing a practitioner’s focus and artifacts for continued reference, which can assist in the follow up of personalized goals. Having success in a personalized goal or experiencing a peer’s success can positively impact teachers’ self-efficacy and their belief in what they can accomplish in the classroom. Video analysis with focused forms, such as rubrics or checklists, gives teachers a tangible way to focus and measure this success.

Recommendations for Future Research

Additional research is needed on how to best structure video analysis, both for self-analysis and for peer review groups in teacher preparation. The nature of feedback that supervisors and mentors provide also offers a rich opportunity for further research. As per Bandura's assertion, "efficacy builders" encourage others' belief in their own ability to perform, but can just as easily undermine self-efficacy with overly negative feedback or setting up too challenging of expectations. Knight (2014) and Marker and D'Onfrio (2010) assert that for video analysis to be effective, there must be a high level of trust in order for video analysis of personal instruction to be a positive experience. Without proper care and expectation-setting, it can easily be a negative and threatening experience. Video recordings also bring up privacy and confidentiality concerns that need to be thoughtfully considered. However, if trust and confidentiality can be appropriately established, there are rich opportunities to apply a microscopic, focused approach to one's classroom interactions. Further research on the kind of mentor and supervisor practices that most develop teacher candidates' self-efficacy would help train those in support roles. Qualitative research would lend insight on students' perceptions and openness to video analysis both as a self- and peer-analysis activity.

Replication studies of this correlational analysis could provide further evidence of these findings across EPPs. It would be valuable to know if other states that use the edTPA find similar results. If continued evidence emerges that writing ability does not predict edTPA performance, that common conception would need to be seriously challenged within EPPs. The evidence from this current study provides rationale to re-think this shared notion among EPPs. A replication study regarding self and peer video

analysis as a predictor of edTPA performance would also be beneficial. Does analyzing others' videos in a video-club-like activity predict edTPA performance, and in particular, does it have a stronger positive relationship with Task 3: *Assessment*? In addition to correlational replication studies, experimental studies with teacher candidates would provide important evidence of cause and effect. A randomized control trial is needed to determine if the presence and type of video analysis activities during the EPP cause improved performance on the edTPA, a proxy for teacher competence and skill.

Conclusion

Video analysis offers the chance for multiple people to give feedback on classroom instruction and to apply multiple focusing questions to the same video (e.g., equity issues, student engagement, conceptual thinking, etc.). Structured video analysis has been shown to engage teachers more deeply in their examination of teaching and learning, rather than overly emphasizing themselves or issues of classroom management when they must rely on memory alone (Knight, 2014; Schieble et al., 2015). Providing evidence of what a teacher is doing well, reinforcing certain strategies, and giving positive feedback can have a greater effect size than corrective feedback (Fukkink et al., 2011). This aligns with Bandura's social learning theory, which suggested that affirmations boost self-efficacy and increase desired behaviors more than critiques. In addition, feedback should be specific to provide information for closing the gap between current and desired performance (Hattie & Timperley, 2007). Video analysis allows for targeting very specific events that can be replayed, slowed down, and discussed with others so that those golden moments are not lost or hidden amidst the flurry of in-the-moment instruction (Knight, 2014; Nagro et al., 2016; Sherin & van Es, 2003). This

practice of targeted video reflection may help teachers develop a habit of mind that transfers to in-the-moment teaching situations, but it requires thoughtful implementation in preparation programs. Future research can add to the growing body of literature that demonstrates how to structure and leverage video analysis for maximum instructional benefit.

References

Adams, J., & Manuel, A. (2016) *Grow your own teachers: Enhancing educator pathways to address teacher shortage and increase diversity.* Professional Educators Standards Board.

American Association of Colleges of Teacher Education. (2017). *edTPA state policy FAQs.* http://edtpa.aacte.org/state-policy

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033295x.84.2.191

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122-147. https://doi.org/10.1037/0003-066x.37.2.122

Bandura, A. (1994). Self-efficacy. In V.S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). Academic Press. https://www.uky.edu/~eushe2/Bandura/BanEncy.html

Beck, R. J., King, A., & Marshall, S. K. (2002). Effects of videocase construction on preservice teachers' observations of teaching. *Journal of Experimental Education, 70*(4), 345-361. https://doi.org/10.1080/00220970209599512

Beisiegel, M., Mitchell, R., & Hill, H. C. (2018). The design of video-based professional development: An exploratory experiment intended to identify effective features. *Journal of Teacher Education, 69*(1), 69-89. https://doi.org/10.1177/0022487117705096

Blomberg, G., Seidel, T., Renkl, A., Sherin, M. G., & Borko, H. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for*

Educational Research Online / Journal Für Bildungsforschung Online, *5*(1), 90-114.

Blomberg, G., Sherin, M., Renkl, A., Glogger, I., & Seidel, T. (2014). Understanding video as a tool for teacher education: Investigating instructional strategies to promote reflection. *Instructional Science*, *42*(3), 443-463. https://doi.org/10.1007/s11251-013-9281-6

Bohan, C. (2016). The past, present, and future of teaching and teacher education curriculum. *Curriculum and Teaching Dialogue, 18*(1&2), 3-12.

Brownell, M. T., Benedict, A. E., Leko, M. M., Peyton, D., Pua, D., Richards-Tutor, C., & Maheady, L. J. (2019). A continuum of pedagogies for preparing teachers to use high-leverage practices. *Remedial & Special Education, 40*(6), 338-355. https://doi.org/10.1177/0741932518824990

Ceven McNally, J. (2015). Learning from one's own teaching: New science teachers analyzing their practice through classroom observation cycles. *Journal of Research in Science Teaching*, *53*(3), 473-501. http://dx.doi.org/10.1002/tea.21253

Cole, L. (2018). Differences in performance on a commercially available competency-based exam between traditionally and alternatively prepared teacher candidates. [Unpublished doctoral dissertation]. Seattle Pacific University.

Danielson, C. (2013). *The framework for teaching evaluation instrument.* The Danielson Group. http://www.loccsd.ca/~div15/wp-content/uploads/2015/09/2013 framework-for-teaching-evaluation-instrument.pdf

Darling-Hammond, L. (2006a). Assessing teacher education: The usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education, 57*(2), 120-138. https://doi.org/10.1177/0022487105283796

Darling-Hammond, L. (2006b). *Powerful teacher education: Lessons from exemplary programs.* Jossey-Bass.

Dawson, P. J., Dawson, K., & Forness, S. (1975). Effect of video feedback on teacher behavior. *Journal of Educational Research, 68,* 197-201. https://doi.org/10.1080/00220671.1975.10884745

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., & Goldman, R. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences, 19*(1), 3-53. https://doi.org/10.1080/10508400903452884

Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics : and sex and drugs and rock 'n' roll.* 4th ed. Los Angeles: Sage.

Fukkink, R. G., Trienekens, N., & Kramer, L. J. C. (2011). Video feedback in education and training: Putting learning in the picture. *Educational Psychology Review, 23*(1), 45-63. https://doi.org/10.1007/s10648-010-9144-5

Gall, M., Gall, J., & Borg, W. (2007). *Educational research: An introduction*. Pearson.

Goldhaber, D., Cowan, J., & Theobald, R. (2017). Evaluating prospective teachers: Testing the predictive validity of the edTPA. *Journal of Teacher Education, 68*(4), 377-393. https://doi.org/10.1177/0022487117702582

Greenblatt, D., & O'Hara, K. (2015). Buyer beware: Lessons learned from edTPA implementation in New York State. *Teacher Education Quarterly, 42*(2), 57-67.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://doi.org/10.3102/003465430298487

Kennedy, M. (2015). Parsing the practice of teaching. *Journal of Teacher Education, 67*(1), 6-17. https://doi.org/10.1177/0022487115614617

Kim, N. (2019). Challenges of teaching and preparing edTPA. *Art Education*, 21-27. https://doi.org/10.1080/00043125.2019.1534439

Kimbrough, S., Davis, J., & Wickersham, L. (2008). Video feedback and semi-structured interviews for reflection among preservice teachers. *Journal of Education and Human Development, 2*(2), 1-12.

Knight, J. (2014). *Focus on teaching: Using video for high-impact instruction.* Corwin.

Marker, E., & D' Onfrio, A. (2010). A different kind of coaching: The professional preparation of graduate level reading specialists combining videocoaching with concurrent feedback. *College Reading Association Yearbook, 31*, 95-112. https://cdn.ymaws.com/aleronline.siteym.com/resource/resmgr/yearbooks/yearbook_volume_31.pdf#page=109

Marzano, R. (2007). *The art and science of teaching. A comprehensive framework for effective instruction.* Association for Supervision and Curriculum Development.

Medina, J. (2008). *Brain rules: 12 principles for surviving and thriving at work, home, and school*. Pear Press.

Mena-Marcos, J., García-Rodríguez, M., & Tillema, H. (2013). Student teacher reflective writing: What does it reveal? *European Journal of Teacher Education, 36*(2), 147-163. https://doi.org/10.1080/02619768.2012.713933

Merriam, S. (2001). *Chapter 1: Andragogy and self-directed learning: Pillars of adult learning theory* in *New directions in adult and continuing education.* Jossey-Bass.

Morin, K. L., Ganz, J. B., Vannest, K. J., Haas, A. N., Nagro, S. A., Peltier, C. J., Fuller, M. C., & Ura, S. K. (2019). A systematic review of single-case research on video analysis as professional development for special educators. *Journal of Special Education*, *53*(1), 3-14. https://doi.org/10.1177/0022466918798361

Nagro, S. A., deBettencourt, L. U., Rosenberg, M. S., Carran, D. T., & Weiss, M. P. (2016). The effects of guided video analysis on teacher candidates' reflective ability and instructional skills. *Teacher Education and Special Education, 40*(1), 7-25. https://doi.org/10.1177/0888406416680469

National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. The National Academies Press. https://doi.org/10.17226/24783

Newton, S. (2010). *Preservice performance assessment and teacher early career effectiveness: Preliminary findings on the performance assessment for California teachers*. Stanford University, Stanford Center for Assessment, Learning, and Equity.

Parkes, K. A., & Powell, S. R. (2015). Is the edTPA the right choice for evaluating teacher readiness? *Arts Education Policy Review*, *116*(2), 103-113. https://doi.org/10.1080/10632913.2014.944964

Pecheone, L., Whittaker, A., & Klesch, H. (2016). *Educative assessment and meaningful support 2015: edTPA administrative report.* Stanford Center for Assessment, Learning and Equity. https://secure.aacte.org/apps/rl/res_get.php?fid=3013&ref=rl

Riley, B. (2020). Drawing on reading science without starting a war. *Educational Leadership*, *77*(5), 16-22.

Santagata, R., & Sandholtz, J. H. (2019). Preservice teachers' mathematics teaching competence: Comparing performance on two measures. *Journal of Teacher Education*, *70*(5), 472-484. https://doi.org/10.1177/0022487117753575

Schieble, M., Vetter, A., & Meacham, M. (2015). A discourse analytic approach to video analysis of teaching: Aligning desired identities with practice. *Journal of Teacher Education*, *66*(3), 245-260. https://doi.org/10.1177/0022487115573264

Seidel, T., Blomberg, G., & Renkl, A. (2013). Instructional strategies for using video in teacher education. *Teaching & Teacher Education, 34*, 56-65. https://doi.org/10.1016/j.tate.2013.03.004

Seidel, T., Sturmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teachers learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching of that of others? *Teaching & Teacher Education, 27*, 259-267. https://doi.org/10.1016/j.tate.2010.08.009

Sherin, M. G., & van Es, E. A. (2003). A new lens on teaching: Learning to notice. *Mathematics Teaching in the Middle School*, *9*(2), 92-95.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20-37. https://doi.org/10.1177/0022487108328155

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics. 6th ed.* Pearson.

van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education, 10*(4), 571-596. Norfolk, VA: Society for Information Technology & Teacher Education. https://www.learntechlib.org/primary/p/9171/

Whittaker, A., Pecheone, R., & Stansbury, K. (2018). Fulfilling our educative mission: A response to edTPA critique. *Education Policy Analysis Archives, 26*(30). http://dx.doi.org/10.14507/epaa.26.3720

Yusuf, M. (2006). Influence of video and audiotapes feedback modes on student teachers' performance. *Malaysian Online Journal of Instructional Technology, 3*(1), 29-35.

Zepeda, S. (2012). *Professional development: What works.* Eye on Education.

Appendix A

Online Supervision Timeline

Online Grad Teacher Ed/Hybrid Supervision Timeline and General Expectations

Observation for candidates enrolled in online programs consists of videos and artifacts, along with feedback and debriefing conferences from field supervisors. Candidates are observed teaching 8-10 times throughout the year one of three ways—live observation, video recording, or live stream. The 8-10 observations can come in any combination of these methods. Observations and feedback will be aligned to the Internship Performance Criteria (IPC). In addition to the lesson observations, additional video journals may explain IPC categories that are not observable within a lesson, such as an introductory context video, classroom assessment strategies, and family engagement strategies. A video journal shows the candidate, and possibly mentor, sharing context, explaining artifacts, and narrating process. After each video submission, the supervisor will mark feedback in GoReact and then debrief with the candidate via a scheduled phone or web conference. Both videos and synchronous conferencing will be marked in the supervisor observation logs. Video observations and journals will be organized in GoReact.
Artifacts such as lesson plans or photos of student work samples may be included in the GoReact submission.
Videoing students requires permission from parents or guardians. Candidates must obtain written permission for each student to be filmed using the SPU consent form in Sharepoint. Any student who does not have permission to be filmed must be positioned outside of the camera's view. Candidates must also agree to policies shown on the Video Policy and Candidate Consent form.
*Suggested due dates for GoReact videos are listed for candidates who are doing entirely video observations. Your specific schedule will be determined between you and your supervisor depending on geography and the number of live visits.

Date	Activity	Description	Location
Summer	Sign Video Policy and Candidate Consent Webform	You will sign a webform in Sharepoint agreeing to responsible use of video during your time in the program	Sharepoint>Teacher Ed>edTPA>Guidelines, Video Consent Forms… *You will sign this form in 6918 Intro to Teaching in the summer
Summer	Familiarize yourself with GoReact; create GoReact login	Watch a screencast of how GoReact works and ensure your device works with GoReact	https://spu.techsmithrelay.com/Y6cf (< 10 minute screencast)

By mid-Sept.	Collect and store video consent forms from students' guardians	Obtain consent forms and learn which students do not have parental permission to be videoed; "Video Policy and Candidate Consent" in Sharepoint should be signed in EDU 6918	Sharepoint>Teacher Ed>edTPA>Guidelines, Video Consent Forms…> "Student Consent to Video Translations" pdf
By end of Sept.	Classroom context introduction	Introduce your supervise to your teaching context (mentor, building, classroom, curriculum)	GoReact activity: Classroom context introduction
First week of Oct.	Lesson Observation #1	Record (or live stream) your first classroom observation for your supervisor and **email your supervisor** once you've submitted it in GoReact	GoReact: Lesson Observation #1
October 30th	**1ST WRITTEN REFLECTION**	Choose one IPC category to reflect on and use classroom evidence and observations to inform your writing	Email/deliver to field supervisor
By mid-Nov.	Lesson Observation #2	Record (or live stream) your second classroom observation	GoReact: Lesson Observation #2
By mid-Dec.	Lesson Observation #3	Record (or live stream) your third classroom observation for your supervisor	GoReact: Lesson Observation #3
By late Jan.	Lesson Observation #4	Record (or live stream)	GoReact: Lesson Observation #4

		classroom observation for your supervisor	
February 1st	**2ND WRITTEN REFLECTION**	Choose one IPC category to reflect on and use classroom evidence and observations to inform your writing	Email/deliver to field supervisor
By early Feb.	Mid-year conference	Mentor, candidate, and supervisor meet to review the IPC and Dispositions Assessment and to set goals for remainder of year	Determined individually (could be web conference using Skype or Zoom, etc.)
By late Feb.	Lesson Observation #5	Record (or live stream) classroom observation for your supervisor	GoReact: Lesson Observation #5
By late March	Lesson Observation #6	Record (or live stream) classroom observation for your supervisor	GoReact: Lesson Observation #6
By late April	Assessment Video Journal	Choose a major summative assessment (e.g. project, essay, or exam) and choose a couple of student work samples with your feedback already on it and describe the process in a video journal. What was the assignment? Scoring criteria? How were the	GoReact activity: Assessment video journal

		students made aware of the criteria? What will students do with your feedback? What would you do again? Differently?	
By early May	Lesson Observation #7	Record (or live stream) classroom observation for your supervisor	GoReact: Lesson Observation #7
By late May	Lesson Observation #8	Record (or live stream) classroom observation for your supervisor	GoReact: Lesson Observation #8
By May 30th	**3rd WRITTEN REFLECTION**	Choose one IPC category to reflect on and use classroom evidence and observations to inform your writing	Email/deliver to field supervisor
Any time (extra)	Lesson Observation #9	Record (or live stream) classroom observation for your supervisor	GoReact: Lesson Observation #9
Any time (extra)	Lesson Observation #10	Record (or live stream) classroom observation for your supervisor	GoReact: Lesson Observation #10
Any time	Families and community video reflection (IPC category 7)	Video yourself discussing your strategies for engaging families and the local community; include relevant artifacts (e.g. email chains,	GoReact: Families and community

		class website, newsletters, notes from attending an event)	
Any time	Professional development video reflection (IPC category 8)	Video yourself discussing your professional development and collaboration within your school and any outside opportunities; include relevant artifacts (e.g. department meeting agenda, workshop notes, collaborative lesson plans, etc.)	GoReact: Professional development
By early June	End-of-year conference	Mentor, candidate, and supervisor meet for final review of the IPC and Dispositions Assessment and check in on goals	Determined individually (could be web conference using Skype or Zoom, etc.)

Appendix B

Reflections for Student Teaching

Reflective writing is an effective way to process feedback from mentor teachers and field supervisors during student teaching. It is also an effective way to evaluate coursework learning and judge its applicability to field experience. Student teachers select topics for reflection in consultation with their mentor teachers and field supervisors.

In general, student teachers will write three reflections during internship: one at the beginning, one in the middle, and one at the end. (See individual program activity timelines for due dates.)

Each reflection should use one of the program standards as an overall theme for writing. In addition, the reflection should include specific information describing how to improve in the area under consideration. For example, a reflection could be written on program standard *7 Families and Community*. The theme of the reflection would be work with families and community. Specifically, the reflection might include setting goals and trying alternative strategies for improving communication with parents about student progress.

Each reflection should state the program standard in the introduction to act as an overall theme for your writing. The content of the reflection should address these questions:

1. **What have you learned about yourself because of your field experience, course content and/or activities?**
2. **How has your learning or perspective changed because of the field experience, course content, and/or activity?**
3. **What are the implications of what you have learned? What will you do differently to more effectively impact your students?**

Student teachers submit their reflections as Word documents to their field supervisors as an email attachment or in person. The length of each reflection should be 600 to 800 words. Writing in APA style is not necessary. Include at least one reference and citation. The reflection should meet requirements for conventions, such as grammar, spelling, and punctuation.

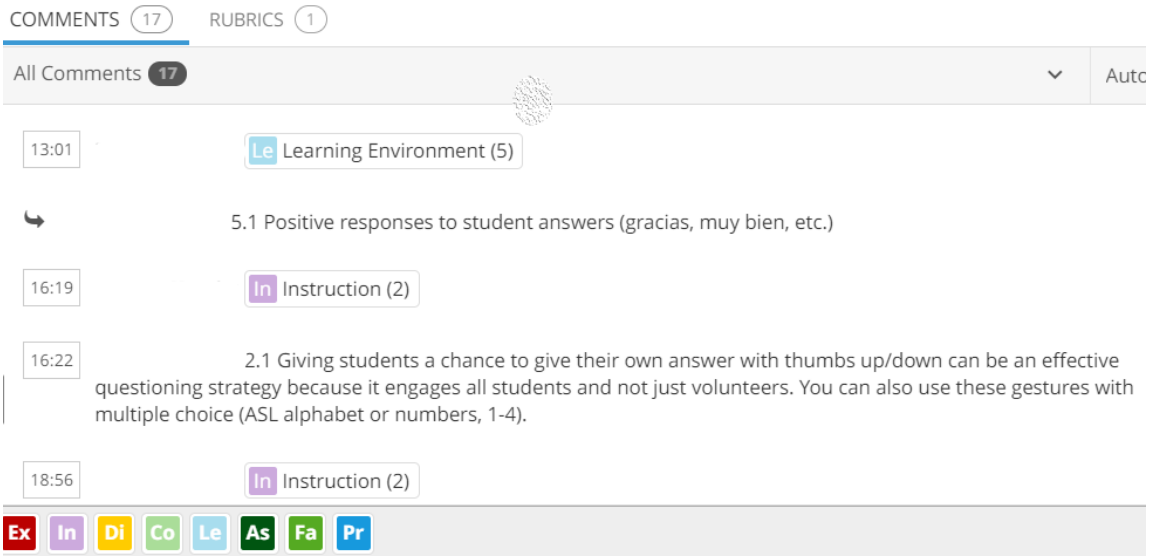
Reflections, which are 10% of grades earned for internship, are scored according to criteria shown below.

Level 1	Level 2	Level 3	Level 4	Level 5
Candidate proposes new insights and changes to teacher practice but provides limited or no analysis of them.	Candidate proposes new insights and changes to teacher practice and **provides analysis of what went well and what improvements need to be made**	Candidate proposes new insights and changes to teacher practice and provides analysis of what went well and what improvements	Candidate proposes new insights and changes to teacher practice and provides analysis of what went well and what improvements	Level 4 plus: **Candidate justifies changes using principles from research and/or theory** Level 4 plus: **Based on the analysis of this self-reflection,**

	to effectively impact students.	need to be made to effectively impact students. **Candidate reflects further on the implications of new learning. Candidate makes superficial connections to research and/or theory.**	need to be made to effectively impact students. Candidate reflects further on the implications of new learning. Candidate makes superficial connections to research and/or theory. **Candidate makes connections to research and/or theory.**	**candidate provides a rationale for • any changes to the strategies proposed OR • why no changes are needed.**

Appendix C

Video Coaching Platform Illustration of Markers and IPC Alignment



Appendix D

Sample Internship Performance Criteria (IPC) Molar Rating Scale

1. Expectations: The teacher communicates high expectations for student learning.

1.1 Component 1a: Demonstrating Knowledge of Content and Pedagogy			
1 Unsatisfactory	**2 Basic**	**3 Proficient**	**4 Distinguished**
Plans and practice display little knowledge of the content, prerequisite relationships between different aspects of the content, or the instructional practices specific to that discipline.	Plans and practice reflect some awareness of the important concepts in the discipline, prerequisite relationships between them, and the instructional practices specific to that discipline.	Plans and practice reflect solid knowledge of the content, prerequisite relationships between important concepts, and the instructional practices specific to that discipline.	Plans and practice reflect extensive knowledge of the content and the structure of the discipline. The teacher actively builds on knowledge of prerequisites and misconceptions when describing instruction or seeking causes for student misunderstanding.
* The educator of students with exceptionalities knows student strengths and challenges when planning accommodations and to ensure that accommodations are implemented across content areas.			
Feedback (strengths, weaknesses, steps for improvement)			

1.2 Component 3a: Communicating with Students			
1 Unsatisfactory	**2 Basic**	**3 Proficient**	**4 Distinguished**
Teacher's explanation of the content is unclear or confusing or uses inappropriate language.	Teacher's explanation of the content is uneven; some is done skillfully, but other portions are difficult to follow.	Teacher's explanation of content is appropriate and connects with students' knowledge and experience.	Teacher's explanation of content is imaginative and connects with students' knowledge and experience. Students contribute to explaining concepts to their peers.
* The educator of students with exceptionalities gives concise and explicit directions along with systematic and explicit instruction to bridge gaps in student knowledge and experiences.			
Feedback (strengths, weaknesses, steps for improvement)			

1.3 Component 3c: Engaging Students in Learning			
1 Unsatisfactory	**2 Basic**	**3 Proficient**	**4 Distinguished**
The lesson has no clearly defined structure, or the pace of the lesson is too slow or rushed, or both.	The lesson has a recognizable structure, although it is not uniformly maintained throughout the lesson. Pacing of the lesson is inconsistent.	The lesson has a clearly defined structure around which the activities are organized. Pacing of the lesson is generally appropriate.	The lesson's structure is highly coherent, allowing for reflection and closure. Pacing of the lesson is appropriate for all students.
* The educator of students with exceptionalities reviews the previous day's lesson, explicitly provides multiple examples and non-examples to increase students understanding of concepts, and allows for individual and group learning.			
Feedback (strengths, weaknesses, steps for improvement)			

Appendix E

Sample Video Analysis Comments

Subject Area	IPC Category	Example Comments in Video Coaching Platform
Elementary math	6) Assessment	*Intern:* Students were called upon to add to definition of array. Class was asked to agree or disagree. (Self) *Supervisor:* Agree or disagree is a great way to hold students accountable to listening and thinking while someone else is speaking.
English language arts (high school)	2) Instruction	*Intern:* We try really hard to have them interact with the texts for maximum understanding and retention and laying out expectations for annotations is one way we do this. (Self)
World language	5) Learning environment	*Intern:* This student is struggling with this class for diverse reasons. When I notice that he could not complete the task, I stay with him, guide him, and even speak some English in order to be very clear about what the task is about. He gets help from his peers and, finally, he can produce an answer, in a cooperative way. I wait for him to respond for as much as needed and help him until he responds correctly. In the end, I celebrate with the rest of the class his success. (Self) *Peer:* Teacher patiently helps student to answer the question in Spanish. (Other) *Another peer:* The students cheer for their classmates. The teacher promotes a respectful classroom environment. (Other)