

Exploring the Psychometrics and the Utility of the ProCAD Instrument

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Lynnette Jane Neu

August 2022

Copyright © 2022 by Lynnette Jane Neu

All Rights Reserved

To my greatest inspirations, my dear nieces,

Mia Grace,

Tess Marie,

Oaklyn Kate,

May this dissertation serve as a reminder that you can complete difficult tasks. When life gets tough, may you not only recall the words of the witty and wise, Lily Tomlin:

“The road to success is always under construction”

...but also remember that Aunt Nettie is only a call away.

Acknowledgements

Thank you to my committee members, Dr. Steve McCallum, Dr. Amelia Brown, Dr. Sherry Bell, and Dr. Merilee McCurdy. Your wisdom, guidance, and feedback have been instrumental in this dissertation's successful completion. Dr. McCallum, thank you for being my committee chair and guiding me through this last year of academic work; your insights were immeasurable. Dr. Brown, I very much so appreciate your mentorship and assistance over the last three years in work, in research, and in life. Thank you also to Dr. Brown and Dr. Judd Laughter for guiding me into the world of teacher dispositions and thoughtfully creating the ProCAD instrument. Dr. Bell, thank you for bringing your unique expertise from both school psychology and teacher education to my committee. Lastly, Dr. McCurdy, thank you for your incredible patience and your bountiful willingness to fulfill various roles for me throughout my graduate school adventure.

To my parents, Dennis and Jane Whitsitt, thank you for believing in me, for accepting me, and for reminding me of my capabilities. I would not be here without you. To my siblings—by blood and by law—thanks for talking my brain down while also lifting my spirits up. Also, thank you for bringing my favorite nieces and furry nephews into my life (and for regularly feeding me). My beautiful cohort, thanks for the shared adventures: Bonfires, talks, and walks. To my roommates, Janet and Kristen, thank you for helping me *reasonably* balance life. To all my friends, new and old, you fill my heart. Grandma, Grandpa, Memom, and Poppa: You are my heroes. Lastly, to my sweet, patient, loving husband, I adore you. Thank you for not only supporting me through all the stressful parts of schooling, but for also doing so amid a global pandemic while also planning our wedding three times, completing tax season from home, and adopting our perfectly particular pup, Tilly. I am so lucky to do life with you, Corbin Neu.

Abstract

In order to evaluate the psychometric properties of a measure of Teacher Candidates' classroom dispositions, the Professional Competencies, Attitudes, and Dispositions (ProCAD), it was administered to 189 yoked rater triads (i.e., Teacher Candidates, Mentor Teachers, and Faculty Supervisors) during the 2018-2019 school year at two timepoints (i.e., Middle & End of professional experience) According to results from exploratory factor analyses, the ProCAD yields one factor. Internal consistency reliabilities for the ProCAD are strong (Range: $\alpha = .88 - .93$). Interrater reliability was assessed through various methods. Two-way, *Proficiency* agreement was measured for each of the eight items and had the highest agreements (Middle & End; range: 70.33% - 98.72%). Results from a mean difference analysis (i.e., Two-Way, Repeated measures ANOVA) revealed statistically significant main effects for time ($p < .001$) and stakeholders ($p < .001$) as well as a significant interaction effect, which became the focus of interpretation as dispositional ratings depend on both stakeholder and time, $F(2, 266) = 15.01, p < 0.001$, partial η^2 [eta – squared] = 0.10. At the Middle point, based on a follow-up ANOVA, stakeholders' scores were significantly different from each other, $F(2, 266) = 27.42, p < 0.001$. Results of post-hoc, pair-wise comparisons showed that Teacher Candidates' ($M = 26.94, SD = 3.26; p < 0.001$) and Mentor Teachers' ($M = 26.88, SD = 3.47; p < 0.001$) ratings were significantly higher than Faculty Supervisors' ratings ($M = 23.57, SD = 4.21$). At the End point, ratings were similar across raters, $F(2, 266) = 0.31, p = ns$. Based on pairwise comparisons with a Bonferroni Correction, dispositional ratings also increased from the Middle of the professional experience to the End for all three raters: Candidates' ($p = 0.048$), Faculty Supervisors' ($p < 0.001$), and Mentor Teachers' ($p = 0.02$). Limitations and implications are discussed.

Table of Contents

Chapter I	1
Literature Review.....	1
A Brief History of Dispositions in Teacher Education.....	4
Operational Definitions of Dispositions	7
Current, Pre-established Measures.....	10
Culturally Responsive Teaching (CRT) Disposition Measures.....	10
Beliefs About Teaching, Second Edition (BATS2).....	11
Watermark-Educational Disposition Assessment (EDA).....	11
Current Programmatic Instruments.....	12
Teacher Disposition Scale (TDS).	12
Teacher education dispositions rating form.....	13
Professional Dispositions Qualities (PDQ).....	14
The Self-Assessment disposition survey.	14
Teacher candidate disposition survey.	15
Candidate Beliefs Self-Assessment Survey (SAS) & Candidate Dispositions Performance Assessment Rubric (CDPA).....	15
Dispositions: Self – and Faculty-Assessment.	16
Summary of Dispositional Measures’ Characteristics and Outcomes.....	17
Instrument Structure.....	17
Changes Over Time.	18
Multi-rater Strategies.	18
The ProCAD Instrument.....	20

Rationale for the Study and Research Questions	21
Chapter II	23
Methods	23
Participants.....	23
Measures	24
Instrumentation	25
Reliability.....	25
Validity	26
Procedures.....	28
Data Analysis	28
Chapter III.....	32
Results.....	32
Data Cleaning.....	32
Missing Variables Analysis	33
Research Question 1: ProCAD Structural Integrity.....	33
Exploratory Factor Analyses.....	34
Research Question 2: Reliability	35
Absolute Agreement	35
Proficiency Agreement	36
Research Question 3: Agreement Across Rater Groups: Intraclass Correlations Coefficients (ICC)	37
Two-Way, Repeated-Measures ANOVA	37

Research Question 4: Do Ratings Differ Across Time?	38
Research Question 5: Do Stakeholders' Mean Scores Differ?	39
Chapter IV.....	41
Discussion.....	41
Research Question 1: ProCAD Structure.....	41
Research Question 2: ProCAD Reliability Estimates	43
Research Question 3: Agreement Across Rater Groups: Interrater Correlations	46
Research Question 4: Stakeholder Differences.....	46
Research Question 5: Changes Over Time	48
Limitations and Implications for Future Research and Practice	50
Summary and Conclusion	56
References.....	57
Appendices.....	70
Appendix A.....	71
Appendix B	81
Appendix C	85
Vita.....	97

List of Tables

Table 1: Means and Standard Deviations for each ProCAD across the three stakeholders and the two time points	72
Table 2: Frequency Counts of Program Areas in the 2018-2019 professional experience cohort with ProCAD ratings	73
Table 3: Frequency Counts of Program Areas in the 2018-2019 professional experience cohort with ProCAD ratings	74
Table 4: Middle of Professional Experience ProCAD Score Agreements	75
Table 5: End of Professional Experience ProCAD Score Agreements	76
Table 6: ProCAD Score Intraclass Correlation Coefficients	77
Table 7: Means and standard deviations for stakeholders' dispositional scores based on time point	78
Table 8: Ranges of Total Averages and Domain Averages with ProCAD Descriptors	79

Chapter I

Introduction and Literature Review

Dispositions define important characteristics of teacher behavior beyond content knowledge and skills (e.g., attitudes toward students, school, and teaching pedagogy, self-reflection, curiosity, open-mindedness), and are considered essential for high-quality instruction. The purpose of this study is to help evaluate the psychometric integrity and utility of an instrument created to operationalize Teacher Candidates' dispositions: The Professional Competencies, Attitudes, and Dispositions Instrument (ProCAD; Brown & Laughter, 2020). The specific goals include determining the structural integrity of the ProCAD instrument, its internal consistency and interrater reliability, the extent to which student candidates' and two supervising mentors' ProCAD ratings converge, and whether ProCAD scores change during a professional experience, presumably as a function of instruction. These goals are consistent with the perspective that teacher performance is best assessed through multiple evaluative channels (Peterson, 2004; Rockoff & Speroni, 2011), and these results should provide a better understanding of the psychometric quality and utility of this dispositional measure.

Literature Review

If education is integral to humankind's improvement, then teachers are society's change agents. Insofar as these statements are true, one avenue to create change is through the programs that teach teachers, known as Educator Preparation Programs (EPPs). The most desirable teachers in the U.S. are trained in teacher preparation programs: the best prepared (Darling-Hammond et al., 2002; Kee, 2012), the most effective (Darling-Hammond et al., 2001), and the longest retained (Fleener & Dahm, 2007; Latham & Vogt, 2007; Redding & Smith, 2016). What

do teachers gain through teacher preparation programs that other avenues to teaching (e.g., transcript review) do not provide?

One answer to this question might be informed by scrutiny of the characteristics of successful teaching, as defined by accrediting bodies such as the Council of Accreditation of Educator Preparation (CAEP), which provide performance standards to EPPs. Typically accrediting bodies divide successful teaching into overarching categories, and CAEP is no exception; specifically, CAEP characterizes teaching performance by three categories: knowledge, skills, and dispositions (CCSSO, 2013). Knowledge and skill assessments are self-explanatory and have been researched extensively, especially after No Child Left Behind highlighted short-term, academic growth as a function of standardized test scores (Hallam, 2009). Oftentimes, students' standardized test scores are used as estimates of teacher performance, but those singular scores do not provide a robust, inclusive portrayal of teachers' performance; hence, data assessing dispositions are also needed (e.g., Boyd et al., 2006; Peterson, 2004; Rockoff & Speroni, 2011). In the sections below, context for the study is provided including a brief literature review defining dispositions, a rationale for the study, and finally, research questions.

Despite six decades of prominence within the education field, it remains unclear whether educators understand and/or appreciate the value of dispositions (Damon, 2016; Harrison et al., 2006; Nweke et al., 2019; Young & Wilkins, 2008). Some of this seeming misunderstanding or ignorance may stem from CAEP requiring dispositional measurement but not providing a straightforward definition (Fonseca-Chacana, 2019). Nevertheless, various invested parties and individuals have developed definitions. Katz and Rath (1985) defined dispositions as the descriptive summation of previous behaviors in specific contexts. Therefore, a disposition is

unique from each behavior determining it. It is the sum of those parts. Council of Chief State School Officers (CCSSO; 2013) described dispositions as critical and as “habits of professional action and moral commitments that underlie the performances [which] play a key role in how teachers do, in fact, act in practice” (p. 6). CAEP (2019) requires that all EPPs candidates demonstrate their understanding of the 10 Interstate Teacher Assessment and Support Consortium (InTASC) standards through their knowledge, skills, and dispositions. The areas within those InTASC standards include the learner and learning; content; instructional practice; and professional responsibility (CAEP, 2019; See Appendix B). In both the 2013 Standards and the 2022 Standards, CAEP only specifies that candidates must “demonstrate positive impact on diverse students’ learning and development” (Council of Accreditation of Educator Preparation [CAEP], 2020b, p. 1). Other specific dispositions are not defined.

Why do dispositions matter? For one, veteran teachers cited characteristics or dispositions (e.g., looking for and accepting change)—as opposed to knowledge or skills—to explain their extended professional retention (Phelps & Benson, 2012). Additionally, comprehensive educator preparation relates to teaching longer (Kee, 2012; Redding & Smith, 2016). Adopting accreditation requirements ensures longer preparation periods, which, presumably, also increases programs’ incentives and time to develop character habits (i.e., dispositions; Brandt et al., 2019). Dispositions are integral for resilient teachers to effectively teach (Darling-Hammond, 2006). Additionally, dispositions are necessary for improving social, emotional, and educational experiences of diverse students in schools (Talbert-Johnson, 2006). Moreover, dispositions that promote productive interactions with students different from oneself (e.g., educationally, racially, ethnically, culturally), can increase through educational teaching experiences (e.g., internship; Frederiksen et al., 2012). In order to have a sufficient number of

teachers for our students, to keep these teachers in the education field, and to maintain understanding teachers who effectively instruct diverse learners, various dispositions needs to be cultivated. Through recruitment and/or thoughtful development, teachers' desirable dispositions could be increased. However, dispositions must be measured *effectively* in order for desirable improvements to be tracked.

A Brief History of Dispositions in Teacher Education

Research regarding teacher dispositions began in 1957 when Arthur W. Combs researched the connection between educators' effectiveness and their attitudes, beliefs, and behaviors (Combs, 1972; Whitsett et al., 2007). Based on theory and research, Combs (1972) concluded that students learn best when instructors engage in the act of becoming teachers—through developing beliefs, characteristics, and attitudes—instead of simply teaching. Teachers who recognize the importance of ensuring students feel safe and understood in order to learn efficiently, regularly act in alignment with such beliefs (Combs, 1972). However, the term “disposition” was seemingly first used in teacher education ten years later. In 1967, Donald Arnstine specifically highlighted teachers' “dispositions” in *Philosophy of Education: Learning and Schooling* (Freeman, 2007). According to Arnstine (1967), dispositions are not directly observable but are a pattern of behaviors combined with attitudes that contribute to traits (Morshead, 1968). Hence, by merely seeing a behavior, it does not follow that a disposition is held (Arnstine, 1967; Morshead, 1968). Moreover, Arnstine (1967) considered teachers' dispositions to be of utmost importance in students' schooling and centered his research around them.

While Combs and Arnstine wrote about dispositions' importance, the teacher education field did not encourage incorporation of dispositions into teacher training until approximately

two decades later (Freeman, 2007). In 1985, Lilian Katz and James Raths defined dispositions as “trends or summaries” of observed actions while also providing non-examples of dispositions (e.g., attitudes, skills; p. 306). Furthermore, they called for EPP’s *employment* of dispositions as part of Teacher Candidates’ competency criterion (Katz & Raths, 1985). Since then, dispositions have become central to teacher education.

Employment of Dispositions (1986-2002). The assumption that dispositions contribute to positive outcomes has become a central theme in educator preparation programs. In 1986, prior to nationwide adoption, the Task Force on Teacher Education for Minnesota’s Future listed dispositions—along with skills and knowledge—as necessary for effective teaching (Freeman, 2007; Minnesota Higher Education Coordinating Board, 1986). Then, InTASC met through CCSSO (InTASC, 1992). InTASC developed ten principles for teachers which were each parsed into knowledge, skills, and dispositions (Freeman, 2007). With the InTASC Standards most recently updated in 2017, dispositions continue to broadly influence education and have been adopted by chief education officers from all states and several additional agencies serving on CCSSO (Council of Chief State School Officers [CCSSO], 2020; See Appendix B). In the decade after 1992, dispositions were also included in the National Council for Accreditation of Teacher Education [NCATE] standards (Freeman, 2007). At that time, multiple shifts occurred simultaneously in teacher education. While many states included a knowledge assessment for teacher licensure to improve teacher quality, in a study of experienced teachers, they highlighted that effective teachers possess more than scholarly expertise (Freeman, 2007).

Recent Disposition Developments (2013-present). In 2010, NCATE and the Teacher Education Accreditation Council (TEAC) announced their merge into a singular accreditation body: CAEP (Leibbrand, 2010). Prior to the merger, NCATE’s most recent standards were

updated in 2008 (National Council for Accreditation of Teacher Education [NCATE], 2008). Dispositions were mentioned throughout their six standards with specification in Standard 1 which was titled “Candidate Knowledge, Skills, and Professional Dispositions” (NCATE, 2008). Additionally, NCATE’s substandards and explanations called for candidate assessments that show growth (NCATE, 2008). However, according to Leibbrand (2010), one of CAEP’s goals was to raise “standards for the evidence which supports claims of quality” (para. 2). This goal was accompanied by specific psychometric requirements for dispositional assessment.

Through its consolidation, CAEP aimed to be the premiere education accreditation body that oversees and improves both candidates’ performance as well as the repository of evidence for the value of its’ standards (Leibbrand, 2010). In 2013, CAEP released their five EPP accreditation standards (CAEP, 2019). Within these standards, CAEP (2019) prescribed new dispositional assessment requirements:

3.3 Educator preparation providers establish and monitor attributes and dispositions beyond academic ability that candidates must demonstrate at admissions and during the program. The provider selects criteria, describes the measures used and evidence of the reliability and validity of those measures, and reports data that show how the academic and non-academic factors predict candidate performance in the program and effective teaching. (p. 2)

While previous NCATE standards assumed credible assessment and data sharing, specific requirements for “reliability and validity” of dispositional measures were new specifications (CAEP, 2019; NCATE, 2008). Hence, many EPPs concluded that previous dispositional assessments (e.g., checklists) were no longer sufficient (Laughter et al., 2020). In addition to

psychometrically sound dispositional assessments, CAEP standards now require that candidates know and understand the ten InTASC standards (CAEP, 2019).

Operational Definitions of Dispositions

Although education dispositional assessments existed prior to CAEP's new standards, most of these measures either lacked reliability and validity and/or alignment with the InTASC standards. Various qualitative methods have been developed to operationalize Teacher Candidate dispositions such as journal entry evaluations (Schussler et al., 2010), reflective essays (Bercaw et al., 2012), and class discussions (Hallam, 2009). Checklists, used by various EPPs, measure a disposition's presence or absence as a function of having observed a student's performance (Beverly et al., 2006; Conderman & Walker, 2015); however, checklists are limited growth measures and do not provide opportunities for constructive feedback for Teacher Candidates (Carroll, 2012; Laughter et al., 2020). Alternatively, many programs have implemented rating scales (Bercaw et al., 2012; Nweke et al., 2019; West et al., 2020; Young & Wilkins, 2008). Rating scales have the potential to be sensitive to multiple raters at multiple times; hence, dispositional growth can be measured more directly.

Due to CAEP's 2013 standards changes, dispositional instrument variability has decreased. However, the flexibility of CAEP's disposition definition means that programs' *original* instruments could and likely do influence their *current* dispositional measurement tools. In 2003, soon after dispositional measurement first became an accreditation standard, Rick Ginsberg and David Whaley (2003) surveyed 27 EPPs with the goal of learning more about their dispositional assessment and their policies tied to these evaluations. At the time, various assessment methods were used to measure candidates' dispositions: disposition checklists (30%), student-faculty conferences (30%), faculty-to-faculty conferences (22%), committee reviews

(19%), and pre-admission letter or contract covering required dispositions (11%; Ginsberg & Whaley, 2003). However, only 76% of the EPPs surveyed had admission and retention policies tied to dispositional assessments (Ginsberg & Whaley, 2003). Therefore, while these programs measured dispositions, there were no formal consequences for failing to meet expectations. Moreover, possibly due to the requirement's newness, the measurement methods were relatively simple and minimally capable of growth measurement.

As disposition requirements matured, the implementation thereof also evolved. In 2008, Young and Wilkins examined the final student teaching assessment for 32 institutions. Their sample was intentionally diverse and representative (i.e., varied geographical location and program size; Young & Wilkins, 2008). An overview of the instruments showed assorted evaluation methods: numeric ratings, narratives, or descriptive target statements (e.g., "proficient;" Young & Wilkins, 2008). Additionally, most instruments included 5-10 dispositional statements (range: 1-25; Young & Wilkins, 2008). The authors defined dispositions as "any specific criteria on the evaluation instrument that did not explicitly refer to knowledge of content or pedagogy or to skills linked to instructional planning, management, delivery, or assessment" (Young & Wilkins, 2008, p. 204). Then, all the dispositional statements were extracted and entered into a database (Young & Wilkins, 2008). Each author coded the statements (e.g., "Reflects on teaching and learning") into pre-defined categories (e.g., "reflects") with an 82% interrater agreement and a qualified outside reader who settled coding discrepancies (Young & Wilkins, 2008, p. 204). Then, the 13 coded categories were sorted into three dispositional themes: (1) Professional behaviors; (2) Professional ways of thinking; (3) Personality characteristics (Young & Wilkins, 2008). Most instruments had a statement coded as 'Establishing relationships with parents and colleagues' ($n = 22$; Young & Wilkins, 2008). The

other most common statements included ‘reflectivity’ ($n = 18$), ‘respect for learners’ ($n = 16$), and ‘work habits and professional growth’ ($n = 15$; Young & Wilkins, 2008).

These studies by Ginsberg and Whaley (2003) as well as Young and Wilkins (2008) were the only two overviews of teacher education dispositions found in the literature. These studies occurred prior to CAEP’s new standards in 2013 and therefore do not necessarily represent the *current* dispositional measurement landscape; however, these articles provide context. Studies of individual programs provide a more robust history of dispositional measurement along with an updated state of affairs.

Old Standards, Old Scales. Prior to CAEP’s 2013 standards, numerous qualitative measurements were used (e.g., Bercaw et al., 2012; Beverly et al., 2006; Conderman & Walker, 2015; Hallam, 2009; Schussler et al., 2010); however, during this time, institutions developed fewer quantitative instruments. Various quantitative instruments that were available measured specific dispositions, especially diversity acceptance and cultural awareness (e.g., Siwatu, 2007). Schulte and others’ (2005) reported that there were no unique instruments that *quantitatively* measured InTASC ’s standards. One year later, although Schulte and others’ *quantitative* measure existed, Flowers (2006) found no *psychometrically-sound*, dispositional measurement instruments. She pointed out that reliability and validity were required for consequential information about Teacher Candidates’ dispositions (Flowers, 2006). Of these very few quantitative rating scales found in the literature published prior to 2013, all had dispositions based on the InTASC standards (Bercaw et al., 2012; Flowers, 2006; Schulte et al., 2005). These instruments had various additional attributes: multiple stakeholder ratings (Bercaw et al., 2012), ratings at various time points (Bercaw, et al., 2012), reliability and validity measures (Flowers, 2006), and statistical confirmations of factorial structure (Flowers, 2006; Shulte et al., 2005).

Current, Pre-established Measures. Using pre-established measures comes with additional work subsequent to instrument selection. If a program has clearly established and operationalized goals, then the use of pre-made tools could be an efficient choice (Immekus, 2016). That is, programs must establish alignment between the instrument and their goals (American Educational Research Association [AERA] et al., 2014). Then, the measure's psychometric properties should be investigated, because it is considered improper to assume that validity and reliability extend beyond the originally measured population (Immekus, 2016). Additional aspects to consider include the instrument's length and cost (Immekus, 2016). Exploring various pre-established instruments and their implementation in programs gives educators a better idea of the pros and cons stemming from their use. Several pre-established instruments are reviewed briefly below: TDI, CRTOE, BATS2, EDA. ***Culturally Responsive Teaching (CRT) Disposition Measures.*** In a California public university's large teacher education program, pre-established surveys gathered information about Culturally Responsive Teaching (CRT) dispositions (Immekus, 2016). Instruments were self-report and included the Teacher Disposition Index (TDI; 45-item), Culturally Responsive Teaching Self-Efficacy Scale (CRTSE; 40-item), and Culturally Responsive Teaching Outcome Expectancy (CRTOE; 26-item) which resulted in high, desirable scores (Immekus, 2016). Teacher candidates tended to rate themselves on the continuum's top half with a median score of four or five, regardless of semester (Immekus, 2016). Due to the data's severe restriction of range, it had limited utility (Immekus, 2016). Additionally, those completing their first semester in the program had similar scores to those preparing to complete the program (Immekus, 2016). As was the case with the TDI, CRTSE, and CRTOE, sometimes pre-established scales' scores may not reflect the real score variance amongst candidates and between time points because of psychometric limitations.

Moreover, pre-existing surveys may only measure a subset (e.g., diversity) of the multiple areas that programs deem important.

Beliefs About Teaching, Second Edition (BATS2). Two institutions' researchers studied their Teacher Candidates' responses to the BATS2 (Lang et al., 2018). This instrument is part of DAATS battery which aims to measure dispositions and also includes the Situational Reflection Assessment (SRA), Candidate Belief Checklist (CBC), Experiences in Teaching Questionnaire (ETQ), and K-12 Dispositions Impact (KIDS; Lang et al., 2018). The BATS2 was developed based on the InTASC standards and the Krathwohl Affective Taxonomy as a framework, with most of the critical dispositions individually included and others being subsumed within additional dispositions (Lang et al., 2018). Lang and others (2018), found that, based on Rasch Analysis, the item scores yielded a normal distribution and supported the instrument's validity; additionally, the instrument had excellent reliability (Lang et al., 2018). Nevertheless, because this instrument is based on self-report, Lang and others suggested that it be used as a screening instrument while utilizing other instruments in the DAATS battery as confirmation of behavior and values.

Watermark-Educational Disposition Assessment (EDA). Possibly the most well-known and "current standard" in Teacher Candidate dispositional measurement comes from the Watermark company for higher education data collection and analysis (2021). It was founded in 2017/2018 through a merger of various major companies in the field: Taskstream, Tk20, and LiveText Digital Measures and EvaluationKIT (Watermark, 2021). For "over seven years," the EDA, LLC researched the Educational Dispositional Assessment's (EDA) components and results (Watermark, 2021). In 2018, EDA, LLC partnered with Watermark. Watermark's platform streamlines education candidates' assessments (Guzzo, 2018). Together, the pair

supports the assessment of teacher dispositions through three instruments, including the EDA (Watermark, 2021). This instrument provides a three-point scoring rubric which can be used at multiple program time points to measure nine dispositions that align with CAEP and InTASC Standards (Watermark, 2021). Through multiple phases of collective agreement, construct validity was reportedly reached; in addition, inter-rater reliability was supported through reportedly “high” correlational coefficients (Watermark, 2021). These reports give weight to pre-established assessments potential to serve as informative and well-aligned dispositional measurement instruments for various education programs.

Current Programmatic Instruments. When CAEP charged EPPs in 2013 to use psychometrically-sound, dispositional instruments, decisions and selections had to be made quickly. Due to various reasons—ranging from high cost to misalignment with goals—many programs decided against using predeveloped instruments and opted instead to develop their own (Immekus, 2016). Moreover, pre-developed options to choose from were limited. Quantitative and psychometrically sound measures aligned with all of the InTASC standards were rare (e.g., Bercaw et al., 2012; Flowers, et al., 2005; Shulte et al., 2005; Siwatu, 2007). In turn, various quantitative instruments were developed with CAEP’s requirements in mind and the freedom to choose specific dispositions aligned with their program’s goals and vision.

Teacher Disposition Scale (TDS). Although not accredited by CAEP, Australia lacked psychometrically-sound dispositional measurement tools. Consequently, West and others (2020) developed the Teacher Disposition Scale (TDS). In order to select dispositions aligned with effective teaching, they searched the literature and then interviewed a panel of Highly Effective Teachers (HAT; West et al., 2020). These HATs were defined and selected by local and federal government (West et al., 2020). Twenty-six final items were selected with a seven-point scale

(West et al., 2020). After that, it was piloted with 179 undergraduate education majors (West et al., 2020). Finally, exploratory factor analysis (EFA) using maximum likelihood estimation and an oblique (direct oblimin) factor rotation was conducted (West et al., 2020). Based on these statistical analyses, 5 subscales (i.e., Motivation to Teach, Teacher Efficacy, Willingness to Learn subscale, Conscientiousness, Interpersonal and Communication) best accounted for variability, and support for this structure was provided through Rasch Analysis (West et al., 2020). Hence, the dispositional scale should not be conceptualized as a singular score as it measures five distinct areas. The TDS exemplifies how a measure with content validity and multiple factors can be developed; however, this article provided no evidence of other psychometric properties such as reliability.

Teacher education dispositions rating form. At a rural, midwestern university, Choi and others (2016) studied whether candidates' disposition ratings related to their students' engagement. The teacher education dispositions' rating form was completed via an online database by field and university supervisors (Choi et al., 2016). Supervisors rated Teacher Candidates' effectiveness on a three-point scale across 19 dispositions (Choi et al., 2016). To gauge student engagement, candidates' classroom students were also observed with the BOSS Modification (Choi et al., 2016). Based on a confirmatory factor analysis (CFA) and subsequent bifactor modeling, the best model included a general dimension along with specific dispositional factors for caring/humanity and responsibility (Choi et al., 2016). While the interrater reliability was low as measured by the correlation between field and university supervisors, there was high internal consistency across the general dimension as measured by omega estimates (Choi et al., 2016). While an underlying factor accounts for much of the test variance, caring/humanity and responsibility are two distinctly different constructs measured by the scale. Although the

interrater agreement was small, the Mentor Teacher and Faculty Supervisor seem to see and to report similarly developed dispositions.

Professional Dispositions Qualities (PDQ). At a large teacher preparation university in the Western United States, Saltis and others (2021) examined the psychometric properties of the Professional Dispositions Qualities (PDQ) survey. This form included twelve dispositions with a three-point Likert-Type rating scale (Saltis et al., 2021). Candidates, Faculty Supervisors, and Mentor Teachers completed it at the program's Beginning, Middle, and End (Saltis et al., 2021). Cronbach's alpha showed high internal consistency ($\alpha = .85$). Not only were Mentor Teachers' ratings significantly higher than Candidates' and Faculties', but generally, ratings increased significantly over time (Saltis et al., 2021). In turn, the scale held together well as a singular dispositional measure, but Mentor Teachers' ratings did not align with Faculty Supervisors (Saltis et al., 2021). This misalignment differs from other instruments. However, the significant time-based improvements support the notion that dispositions can and do change through experiences.

The Self-Assessment disposition survey. At a regional college in Southeastern Nebraska, in a dissertation, Seay (2017) explored changes in Teacher Candidates' dispositional self-evaluations over time as measured by the institutionally-developed, dispositional instrument. Three statements aligned with the college's various accrediting bodies' disposition standards and were included in the analysis (Seay, 2017). Candidates self-evaluated using a four-point scale at the Beginning and the End of the teacher education program (Seay, 2017). Through an ANOVA, Seay (2017) found that dispositions increased significantly from the Beginning to the End of the

program. These results suggest not only that dispositions can change, but also that these changes can be measured with institutions' instruments.

Teacher candidate disposition survey. At a large university, Henry and others (2013) used five measures of Teacher Candidates' performance in the program (i.e., coursework and grades, professional behaviors and dispositions, student teaching performance assessment, Praxis I exam scores, and comprehensive portfolios) to predict student learning (i.e., math achievement scores, reading achievement scores) in grades 3-5. Early in preparation, the Teacher Candidate completed the dispositional survey, and then their Mentor Teacher completed it towards the end of student teaching (Henry et al., 2013). Based on an exploratory factor analysis, the university's dispositional instrument with 20 items can be best represented by a single underlying factor as opposed to three to four underlying factors, as anticipated (Henry et al., 2013). Additionally, dispositional scores did not predict grade 3-5 student achievement (Henry et al., 2013). Despite structural expectations, this dispositional measure is best characterized by a general underlying dimension and does not directly predict future students' achievement.

Candidate Beliefs Self-Assessment Survey (SAS) & Candidate Dispositions

Performance Assessment Rubric (CDPA). Four universities (2 private and 2 public; 1 historically Black college/university; 1 religiously-affiliated university) were included in this study by Nweke and others (2019). These dispositional measures, Candidate Beliefs Self-Assessment Survey (SAS) & Candidate Dispositions Performance Assessment Rubric (CDPA), were developed through examination of previous research with K-12 and higher education professionals (Nweke et al., 2019). Twenty-four dispositions were chosen that had appropriate content validity ratios (CVR; Nweke et al., 2019). Then, the development team created indicators for the CDPA, belief statements for conceptualization, and self-assessment statements for the

SAS (Nweke et al., 2019). They created a 4-point scale for the CDPA (Nweke et al., 2019). A 5-point scale was used for the SAS which included an “undecided” option because it was administered at program entry when candidates may not understand the dispositions, and the “undecided” option may be their most appropriate selection (Nweke et al., 2019). Both instruments yielded high reliability coefficients, based on Cronbach’s alpha (CDPA alpha = 0.96; SAS alpha = .81). Lastly, for the CDPA, intraclass correlation coefficients (ICC) were calculated to determine interrater reliability (Nweke et al., 2019). Based on averages, the ICC was .80 (95CI = .67-.90). These findings illustrate how reliable instruments can be when used by various parties at different time points. More specifically, according to Koo and Li (2016), the ICC of .80 is considered a “good” reliability.

Dispositions: Self – and Faculty-Assessment. At a large midwestern university, elementary education and special education candidates' dispositions were measured (Conderman & Walker, 2015). While the self- and teacher-ratings were measured with unique instruments, these different instruments had similar items (Conderman & Walker, 2015). The faculty's instrument includes 23 items that were related to eight general indicators with 4-point scales (Conderman & Walker, 2015). Candidates' forms were condensed versions of faculty's scales with fewer items and dichotomous ratings (+/-)(Conderman & Walker, 2015). Candidates completed their forms after a dispositional workshop, which they attended during their first student teaching semester (Conderman & Walker, 2015). Candidates' score reliability was high with a Kuder-Richardson 20 value of 0.88 (Conderman & Walker, 2015). For elementary education, candidates and faculty agreed that areas needing the most improvement included Tardiness, Lack of Study Skills, and Poor-Quality of Written Work (Conderman & Walker, 2015). For special education, the areas of Tardiness and Poor-Quality of Written Work were

agreed upon as needing improvement (Conderman & Walker, 2015). Through these findings, insights into the similarities between faculty and candidates' ratings can be drawn. However, additional studies are necessary before conclusions are drawn due to the limited nature of the population (e.g., program areas), the differences between measurements (e.g., scale; number of items; wording), and the limited nature of the candidates' scale.

Summary of Dispositional Measures' Characteristics and Outcomes. Various measures exist in the field including those that are pre-established and widely distributed (e.g., Immekus, 2016; Lang et al., 2018; Watermark, 2021) as well as those that were developed by particular institutions which have potential for further distribution (e.g., Choi et al., 2016; Conderman & Walker, 2015; Henry et al., 2013; Nweke et al., 2019; Saltis et al., 2021; Seay, 2017; West et al., 2020). While these instruments are different, some underlying similarities link them as well: Structure, Raters, and Changes Over Time. A better understanding of these instruments guided the current study.

Instrument Structure. The instruments described above are tools that were either found in academic literature and/or are well-known in the field. As of fall 2020, CAEP accredited 366 EPPs (CAEP, 2020a); however, only approximately 10 modern instruments were reviewed in this study. While some of these, especially those classified as "pre-established," are likely used by multiple institutions, there are likely many more tools in use. Nevertheless, within this sample, psychometric analyses defined various instrument structures. The underlying factors ranged from a single general factor (e.g., Henry et al., 2013) to as many as five factors (e.g., West et al., 2020). Although all these instruments were designed to assess dispositions, these instruments' structures differ greatly. According to Fonseca-Chacana (2019) the field offers the

freedom to define dispositions differently and to create measurement tools with various properties/structures.

Changes Over Time. Many instruments have been created to assess dispositions at multiple time points (Bercaw, et al., 2012; Immekus, 2016; Saltis et al., 2021; Seay, 2017; Watermark, 2021). Some, such as the TDI, CRTSE, and CRTOE studied by Immekus (2016), do not show dispositional change from early in the program to the End. On one hand, perhaps this finding provides evidence for the theoretical position that dispositions are deeply entrenched and unchanging (e.g., McKnight, 2004; Wasicsko et al., 2009). On the other hand, some, such as Immekus (2016), argue that current instruments are poorly-suited for measuring Teacher Candidate dispositions. However, some instruments (e.g., PDQ; The Self-Assessment Disposition Survey) do show significant changes over a candidate's time in the EPP (Saltis et al., 2021; Seay, 2017). Consequently, adding weight to the theory that dispositions can change (Caroll, 2012; Diez, 2007). Therefore, the extent to which instruments provide evidence of dispositional change, or the lack thereof, illustrates the importance of not only the psychometric integrity of instruments but is also affected by the constructs operationalized by the instruments' creators.

Multi-rater Strategies. Although objective tests have been considered the educational gold standard for decades, recent perspectives acknowledge that no single measure operationalizes truth with certainty (Rockoff & Speroni, 2011; Shepard, 2016). In business contexts, 360-degree feedback (e.g., supervisor, peer, and subordinate ratings) adds to the psychometric integrity of workplace performance measures (Atwater & Brett, 2002; Maurer et al., 2002). Similar to the business environment, the teacher education world includes multiple stakeholders: peer teachers, Mentor Teachers, Faculty Supervisors, and Teacher Candidates that

can offer unique perspectives (Darling-Hammond, 2000). Multi-rater assessments are regularly used in teacher education due to the many parties invested in teachers' development (e.g., lesson plan feedback; Tillema, 2009). More specifically, multiple ratings are often used in candidates' disposition assessment (Bercaw et al., 2012; Beverly et al., 2006; Conderman & Walker, 2015; Nweke et al., 2019; West et al., 2020).

Through a meta-analysis of 115 independent business-based samples, Heidemeier and Moser (2009) found that self-raters tend to be more lenient than supervisory raters. Hence, employees generally rated their own performance higher than their supervisors did, seemingly because employees are more lenient with themselves. Importantly, employees' education level decreased rating discrepancy: employee's self-ratings decreased and became more similar to supervisors' ratings with time and education (Heidemeier & Moser, 2009). Regarding Teacher Candidate dispositions, Conderman and Walker (2015) found that candidates and faculty instructors agreed overall in their dispositional assessments but there were some rating discrepancies regarding specific dispositions (i.e., attentiveness in class). Saltis and others (2021) found that field supervisors rated candidates significantly higher than Candidates and Faculty Supervisors rated them. Nevertheless, through descriptive analysis of a dispositional assessment, Laughter et al. (2020) found that candidates tended to rate themselves higher than supervisors and Mentor Teachers. Additionally, Choi and others (2016) found that field supervisors' and Faculty Supervisors' ratings were only minimally correlated. Hence, within the education field, comparisons between supervisors and supervisees should be the focus of additional research. These and related test development issues have been the focus of those who developed the ProCAD instrument.

The ProCAD Instrument

Similar to other EPPs, the University of Tennessee-Knoxville created its own instrument to measure dispositions with a foundation of CAEP's guidelines and the InTASC standards. Each InTASC standard is broken into three subcategories: performances, knowledge, and dispositions (CCSSO, 2013). The University of Tennessee-Knoxville teacher preparation program determined that the first two standards were operationalized adequately by existing instruments—performance (e.g., observations) and knowledge (e.g., GPA, Praxis Scores), but that the dispositional checklist originally developed to operationalize dispositions was insufficient based on CAEP's 2013 standards (Laughter et al., 2020). Consequently, the Professional Competencies, Attitudes, and Dispositions (ProCAD) instrument was developed (see Appendix C).

ProCAD is used at various times by multiple stakeholders: Faculty Supervisors, school mentors, and Teacher Candidates (Laughter et al., 2020). Each stakeholder rates the Teacher Candidate on a scale: (1) Unsatisfactory, (2) Developing, (3) Proficient, and (4) Exemplary (Laughter et al., 2020). In turn, this system can measure candidate growth from program acceptance to professional experience completion. While an (1) Unsatisfactory score always generates an Area of Refinement Plan (ARP), a score of (2) Developing is considered an appropriate rating, especially early in the program, as Proficiency is not expected until professional experience completion (Laughter et al., 2020). Through Teacher Candidate self-assessment paired with outside ratings, candidates are given various reflection opportunities, which is also a CAEP standard (CAEP, 2019). Reflective teaching and teacher retention are positively related (Kelley, 2004); thus, an additional benefit of the ProCAD instrument is that it

compels Teacher Candidates to practice self-reflection. Development of the ProCAD is further described in the Method section as it is the primary focus of this study.

Rationale for the Study and Research Questions

The UT teacher preparation faculty projects using ProCAD for the foreseeable future and anticipates collecting data over multiple semesters as well as at least one CAEP accreditation cycle in the future years. Researching this instrument's psychometric characteristics and capacity to inform both candidates and supervising mentors regarding candidate performance can enhance stakeholders' understanding of the instrument and how it can inform candidate progress and operationalize program changes/improvement. While Laughter and others (2020) provided important, initial information about the ProCAD instrument through descriptive statistics, use of analyses requiring descriptive and inferential statistics will augment evidence of the instrument's psychometric integrity and utility. As outlined by Tillema (2009) instrument utility is informed by knowledge of item characteristics and related psychometric properties, evidence of rater sensitivity, and ability to inform stakeholders about candidate progress. Educators need empirical data addressing these issues. Consequently, the goal of the current study is to provide data informing the psychometrics of the ProCAD instrument and the extent to which scores differ as a function of rater perspective and time. The following research questions will be addressed:

1. To what extent is the factor structure of the ProCAD equal across candidate, mentor and supervisor (e.g., to what extent do the eight individual rubrics/items scores on the ProCAD instrument comprise a similar number of factors/scales across raters?)
2. Is there evidence for ProCAD internal consistency reliability determined from Cronbach's Alphas and interrater reliability as determined by percentage agreement

- based on absolute and proficiency agreement across three raters (i.e., Teacher Candidate, Mentor Teacher, and Faculty Supervisor)?
3. What are the relationships among the three stakeholders scores (i.e., Teacher Candidate, Mentor Teacher, and Faculty Supervisor), as determined by group agreement of variance estimates using an intraclass correlational analyses (ICC)?
 4. Do the three stakeholders' mean scores differ significantly at two times during the internship (i.e., Middle and End points)?
 5. Do dispositional scores change significantly from the Middle of the professional experience to the End as determined by mean gain scores?

Chapter II

Methods

Participants

Based on the presence of a rating by at least one stakeholder during at least one time point, ProCAD data were analyzed for 189 Teacher Candidates from the 2018-2019 school year. At this Research-Intensive university in the Southeast, most of the program's teacher licensure areas were represented by this sample of Teacher Candidates (see Table 2). Initially, candidates submitted 313 self-ratings. Two-hundred, seventy-six Mentor Teacher assessments and 208 Faculty Supervisor assessments were also obtained for candidates. These total counts include data submitted at three times during candidates' professional experience year: Beginning (initial), Middle, and End. Within these initial numbers, three professional experiences offered by the university were represented: Internship ($n=116$, 61.4%), Student Teaching ($n=53$, 28.0%), and Job-Embedded Practitioner ($n=20$; 10.6%). For various analyses, cases were excluded pairwise; therefore, the sample number also varied.

Demographic data and test scores are not included in the ProCAD collection process nor was this information requested for the study. However, this information was reported in aggregated form on the TN EPP Report card (Tennessee State Board of Education, 2019). During 2018-2019, the percentage of racially diverse cohort members (8.70%) fell below the state average (15.6%): Asian (2.1%), Black (2.5%), White (91.3%), and Two/More Races (2.1%). One-hundred percent of students earned qualifying assessment scores which was higher than the state average ($M = 94.0\%$). Additionally, 28.9% of student earned licensure in high-demand endorsement areas which was also higher than the state average ($M = 27.8$).

Measures

Interrater reliability estimates (i.e., *Absolute* agreement, *Proficiency* Agreement) and stakeholders' ratings across ProCAD domains were obtained for the following domains. ProCAD contains eight items, each one designed to assess one of the following domains:

1. Attitudes towards Teaching Methods
2. Attitudes toward Students
3. Attitudes toward Schools
4. Open-Mindedness Dispositions
5. Self-Reflection Dispositions
6. Curiosity Dispositions
7. Educational Equity Dispositions
8. Professional Competencies

All three stakeholders (i.e., candidates, Faculty Supervisor, Mentor Teacher) rated candidates on a scale (1 to 4) at three time points during their professional experience year (i.e., Beginning, Middle, End; Laughter et al., 2020). Each of these domain items uses the following rating scheme; (1) unsatisfactory, (2) developing, (3) proficient, (4) exemplary. Each rubric also includes a general explanation of what the domains measure along with the InTASC sub-standards that underlie the domain (see Appendix C).

Both *Absolute* agreement and *Proficiency* agreement were calculated based on percentages for each of the eight ProCAD domains. *Absolute* Agreement is defined as the percentage of agreement between stakeholders' exact ratings. Therefore, for example, if, on a particular domain, a Faculty Supervisor rates the candidate as a "3" and the Mentor Teacher rates the candidate as a "3" this agreement is considered Absolute; on the other hand, if the Faculty

Supervisor rates the candidate as a “3” and the Mentor Teacher rates a candidate as a “4,” the two do not agree *Absolutely*. Most importantly, the definition of *Proficiency Agreement* is the agreement between stakeholders’ ratings of proficiency (Laughter et al., 2020); ratings of “3” and “4” are considered *proficient*. Rankings of “1” and “2” are considered *not* proficient (Laughter et al., 2020). Therefore, a Faculty Supervisor’s rating of “3” would agree with a Mentor Teacher’s rating of “4” based on *Proficiency* agreement. Both *Proficiency* and *Absolute* agreements will be calculated across each pair of stakeholders (i.e., Teacher Candidates & Faculty Supervisors, Teacher Candidates & Mentor Teachers, Faculty Supervisors & Mentor Teachers) for “Two-Rater Agreement” as well as across all three stakeholders for “Three-Rater Agreement” (Laughter et al., 2020). Only the two stakeholders being compared need to have the same ratings for the Two-Rater comparison to be considered an “agreement.” However, all three stakeholders must use the same rating for the Three-Rater comparison to be classified as an “agreement.”

Instrumentation

CAEP has specific expectations of instrument validity and reliability and there are multiple best practices in the field of teacher education. Based on these requirements and expectations, various forms of validity (i.e., face, content, predictive) and reliability (i.e., test-retest, interrater) have been calculated and/or determined for the ProCAD instrument.

Reliability

Interrater reliability was established prior to the 2017-2018 school year through multiple methods. Test-retest reliability was also calculated. For a single classroom of an individual course ($n = 18$), the interrater reliability between the course’s graduate assistant and the professor was calculated (Laughter et al., 2020). Actual agreement ranged from 27% to 100%, and

proficiency agreement (i.e., 1s & 2s coded as “not proficient;” 3s & 4s coded as “proficient”) ranged from 94% to 100% (Laughter et al., 2020). Ranges were based on mean agreement percentages that were calculated for each of the five ProCAD areas measured (i.e., Attitudes toward Teaching Methods, Open-Mindedness Dispositions, Self-Reflection Dispositions, Curiosity Dispositions, and Professional Competencies; Laughter et al., 2020). When Teacher Candidates’ self-ratings were included, the three-rater actual agreement ranged from 40.70% to 85.20% and the proficiency agreement ranged from 96% to 100% (Laughter et al., 2020). With a larger group of candidates ($n=127$), Faculty Supervisors, and Mentor Teachers, the three-rater actual agreement decreased to a range of 43.7% to 56.8% and the proficiency agreement ranged from 81.0% to 91.9% across the eight domains (Laughter et al., 2020).

Test-retest reliability was measured with Teacher Candidates’ self-ratings. Candidates attended a ProCAD training. Then, they completed the ProCAD assessment both before and after a few hours orientation that covered topics unrelated to ProCAD (Laughter et al., 2020). For all eight domains, *Absolute* agreements exceeded 73% and *Proficiency* agreements exceeded 90% (Laughter et al., 2020). This ProCAD assessment is designed to test reliability and validity over time through continuing interrater agreement measures as well as content evaluation through open-ended comments (Laughter et al., 2020).

Validity

To establish face, criterion, and content validity, the ProCAD instrument was developed based on the professionally established InTASC standards (Laughter et al., 2020). Through discussion and Lawshe content validity rankings, UTK’s EPP faculty determined that all the standards were critical and, in turn, were included in the ProCAD instrument (Laughter et al., 2020). Then, all 43 InTASC dispositional substandards were sorted into one of three overarching

categories: attitudes, dispositions, and competencies (Laughter et al., 2020). InTASC and CAEP defined all of these areas as dispositions. However, the university's ProCAD development team decided that the areas were more appropriately defined as a set of beliefs that predispose action (attitudes), daily teacher actions outside of pedagogical necessity (professional competencies), and habitual behaviors in specific contexts (dispositions; Laughter et al., 2020). Through content analysis, the substandards were sorted into one of the three overarching categories: one competency category (i.e., professional competencies), three attitudes (i.e., attitudes towards teaching, students, and schools), and four dispositions (i.e., open-mindedness, self-reflection, curiosity, educational equity; Laughter et al., 2020). The 43 InTASC substandards were included within categories to inform scoring criteria for each rubric (Laughter et al., 2020). Content experts developed initial rubrics which were subsequently reviewed by the standards-based assessment panel (Laughter et al., 2020). Consensus was reached between the two parties and the rubric was finalized (Laughter et al., 2020). The InTASC standards included in each ProCAD domain are listed in the ProCAD rubric (see Appendix C).

A subsequent study of the 2017-2018 internship data ($n = 785$) evaluated some psychometric properties of ProCAD. Significant correlation coefficients were obtained between individual ProCAD rubric scores and several other measures of candidates' success (Laughter et al., 2020). These relationships address the predictive or criterion validity of this measure. Measures with lower correlations included the ACT and undergraduate GPA scores (Laughter et al., 2020); higher correlations were found between ProCAD and edTPA as well as university graduate-level coursework (Laughter et al., 2020). EPPs across the nation use edTPA as a support and an assessment instrument in 27 subject areas to "emphasize, measure and support [candidates'] skills and knowledge" required for classroom success (edTPA, 2021, para. 3).

Because ProCAD, edTPA, and coursework are all related to and based on the InTASC standards, these significant correlations were expected (Laughter et al., 2020). These results add to the data base in support of ProCAD's content and criterion validity and show the nature of important relationships between dispositions and other educationally relevant variables. For example, ProCAD's dispositions are somewhat independent of other measures, such as ACT's measurement of skills (Laughter et al., 2020).

Procedures

Data included in this study are considered "institutional data," which means that the information was collected for purposes other than this study itself. In order to use this data, Institutional Review Board (IRB) approval was gained. Subsequently, the approved procedures were utilized.

For institutional purposes, these data are collected at various time points throughout Teacher Candidates' progression through the program (Laughter et al., 2020). More specifically, the data included in this study were collected during the fall of 2018, winter of 2018, and spring of 2019 through a secure, online data collection form. Teacher Candidates, Faculty Supervisors, and Mentor Teachers accessed the form independently to complete the rating scale at the specified time points.

Data Analysis

First, the nature of missing data was determined. As Henry et al. (2013) suggested, the nature of data gathering in teacher education at various time points from various stakeholders results in many opportunities for data to be missing, as a function of lack of completion, inability to match candidates (e.g., different last name, misspellings), and/or other influences. Little's (1988) Missing Completely at Random (MCAR) test was conducted in SPSS through the

Missing Values Analysis (MVA) in order to determine whether the cases with missing data could be excluded from analyses without influencing the findings' validity.

Next, the structural integrity of ProCAD was determined. Raters' scores from the Middle of the professional experience assessment were included in an exploratory factor analysis. The first unrotated factor structure was obtained initially and evaluated; follow up analyses using Promax rotations were conducted to maximize congruence of scores within factors (Pett et al., 2003). This analysis helped determine the factor structure for each stakeholder.

In addition, reliability analyses were conducted. A single Cronbach's Alpha was obtained for each stakeholder's total score. Also, both the previous methods cited in Laughter and others (2020) of *Absolute* and *Proficiency* agreements were calculated. Neither the current nor the 2022 CAEP standards specify reliability methods. Instead, the current standards require that "The provider selects criteria, describes the measures used and evidence of the reliability and validity of those measures..." (CAEP, 2019, p.2). However, fewer specifics are provided in the forthcoming standards. That is, for 2022, the EPP "creates and monitors transition points from admission through completion" for various aspects of effective teaching including dispositions, but no specifications for psychometric properties or reliability are mentioned (CAEP, 2020b). Consequently, to be consistent with previous methodology exploring the psychometrics of the ProCAD, both *Absolute* agreement and *Proficiency* agreement were calculated based on percentages for each of the eight ProCAD domains. Using methodology from Laughter and others (2020) *Absolute* Agreement is defined as the percentage of agreement between stakeholders' exact ratings; *Proficiency* Agreement is defined as the percentage of agreement between stakeholders' ratings of proficiency. Rankings of "3" and "4" are considered *Proficient*, but rankings of "1" and "2" are considered *not* proficient (Laughter et al., 2020). Moreover, both

Proficiency and *Absolute* agreements were calculated across each pair of stakeholders (i.e., Teacher Candidates & Faculty Supervisors, Teacher Candidates & Mentor Teachers, Faculty Supervisors & Mentor Teachers) for “Two-Rater Agreement” as well as across all three stakeholders for “Three-Rater Agreement” (Laughter et al., 2020). The number of true agreements were divided by the total number of pairs or triads within that comparison to determine the agreement percentage.

Due to the grouping of raters into three groups, Intraclass Correlation Coefficients (ICC) were calculated (Landers, 2015). Since the *same* raters *did not* provide scores for each candidate, the “One-Way Random” ICC is appropriate (Landers, 2015). Moreover, since absolute agreement between the three-raters was already calculated and the “true” dispositional score was unknown, the ICC was reported based on the average of the measures (Koo & Li, 2015). Additionally, these were calculated for each of the eight dispositional domains at the Middle timepoint. The ProCAD instrument is intentionally designed for this kind of continuous reliability calculation (Laughter et al., 2020).

Finally, the overall ProCAD mean scores (dependent variable) were compared across stakeholders (mentor, supervisor, candidate) and time (beginning vs. End of professional experience). Therefore, a repeated measures ANOVA was conducted with post hoc analyses as needed, with two independent variables—Stakeholder (Supervisor, Mentor, Candidate) and Time (Middle of Professional experience vs. End of Professional experience; dispositions were analyzed as a repeated measure). Initially, the intent was to provide an analysis of mean differences across three time points, but the sample size for the first of the three was too small. Interaction effects and main effects are reported. Appropriate simple effects analyses were also conducted to determine the effects of time on scores from Teacher Candidates, Faculty

Supervisors, and Mentor Teachers (Morrow, 2018). Additionally, simple effects analyses were conducted to determine the effects of stakeholder at the Middle and at the End of the professional experience (Morrow, 2018). Effect sizes were determined through partial eta-squared which measure the proportion of variance in the dependent variable associated with changes in the independent variable, relative to error variance (Tabachnick & Fidell, 2007).

Chapter III

Results

In order to assess the psychometric properties and the utility of the ProCAD instrument, the dispositional ratings for 189 Teacher Candidates during the 2018-2019 school year were analyzed. Descriptions of data cleaning, assumption testing, and analyses are described. Specifically, the results of the EFA, interrater agreement, ICC, and two-way, repeated measures ANOVA are detailed. The means, standard deviations, and mean percentage increases for each of the eight domains across the three stakeholders and the two time points can be found in Table 1. Percentage increases ranged from -1% to 17%. The largest percentage increase was for Faculty Supervisors' ratings of the curiosity disposition (17% increase). Teacher candidates' ratings of their self-reflection disposition actually decreased slightly (1%) from the Middle to the End. All other ratings increased from the Middle to the End point (See Table 1). All tables and figures can be found in Appendix A.

Based on histogram inspection, Mentor Teachers' End scores were slightly negatively skewed, but all others were approximately normal. However, parametric tests are robust with respect to violations of normality assuming an appropriately large sample size ($n > 30$), (Tabachnick & Fidell, 2007); consequently, parametric statistical tests were conducted as appropriate (e.g., ANOVAs).

Data Cleaning

The ProCAD instrument was designed to be completed at three time points during the professional experience: Beginning, Middle, and End. However, instead of selecting one of these three options, some applicants wrote in the time point. Due to these variations, all ratings were classified as beginning, Middle, or End based on observation date: August - October

(Beginning); November - March (Middle); April - July (End). November was included in the Middle as opposed to the Beginning due to the extended nature of the winter break in December. Hence, if the scale was completed prior to the university's final examinations and winter break, then it was likely completed in November. These classifications were implemented and maintained throughout data analysis. Data were matched across raters for candidates and all identifying information was stripped from the database prior to the database ownership being transferred to the research team. Procedures outlined in the approved IRB application were utilized.

Missing Variables Analysis

All data analyses, other than *Proficiency* and *Absolute* agreements, were conducted with SPSS® 27. First, a missing variable analysis (MVA) was conducted on all variables. There was a large percentage of missing data at the Beginning of professional experience (87.3%). When the beginning ratings were removed, Little's MCAR test was no longer significant for either of the two methods of data analyses, $\chi^2(18) = 22.69, p = 0.20$; $\chi^2(855) = 398.59, p = 1.00$. That is, the data can be interpreted as missing completely at random and multiple imputation is unnecessary (Garson, 2015). Based on these analyses and the percentage of missing data, the beginning scores were not included in further analyses, because including such a small percentage of beginning data would likely create a source of bias.

Research Question 1: ProCAD Structural Integrity

Prior to additional analyses, Research Question 1 was addressed. As previously noted, the factor structure of the eight ProCAD items for each stakeholder was examined; but first statistics related to the assumptions were calculated: Correlations, Kaiser-Meyer-Olkin measure of sampling adequacy, and Bartlett's test of sphericity. First, for all three stakeholders, all eight

items show evidence of shared variance based on a correlation matrix (i.e., coefficients' range: .36 - .74). Next, the Kaiser-Meyer-Olkin statistic reflecting degree of sampling adequacy as measured by shared variance was above the commonly recommended value of 0.60 for Teacher Candidates (.91), Faculty Supervisors (.90), and Mentor Teachers (.92), reflecting adequate size (Tabachnick & Fidell, 2007). Another operationalization of shared variation was obtained, the Bartlett's Chi-Square test of sphericity, and was also significant for Teacher Candidates ($\chi^2(28) = 499.469, p < .001$), Faculty Supervisors ($\chi^2(29) = 741.434, p < .001$), and Mentor Teachers ($\chi^2(28) = 522.486, p < .001$), adding confidence to the factorability of these domains. Based on these aforementioned criteria, factor analysis was considered appropriate for all three stakeholders across the eight ProCAD items.

Exploratory Factor Analyses

Since the purpose of this analysis was to examine (and maximize) congruence among items, an exploratory factor analysis was conducted using the Oblimin solution with Promax rotation. Based on examination of scree plots and eigenvalues a one-factor solution represented the best fit to the data for all three stakeholders. That is, only one factor yielded an eigen value greater than one for each. All three stakeholders' item communalities were above .3 (range: .42 – .78), reflecting appreciable shared variance across all items. The first factor explained 54.49%, 65.10%, and 66.62% of variance for the Teacher Candidates, Faculty Supervisors, and Mentor Teachers, respectively. All items were retained based on sufficient loadings across all three factors for each rater group (i.e., $>.35$; range: .64 - .88; Tabachnick & Fidell, 2007). See Table 3 for factor loadings for each of the stakeholders.

Results from the exploratory factor analyses informed subsequent analyses. That is, the total ProCAD score was calculated and used for subsequent inferential analyses for all

stakeholders (i.e., ANOVA), even though the instrument was conceptually created by including items across multiple dispositional foci.

Research Question 2: Reliability

As noted above, one general ProCAD factor was retained; consequently, composite scores were created from the eight items for each stakeholder based on the total of the scores, with higher total scores indicative of more desirable dispositions. Internal consistency reliabilities for the composite ProCAD scores were estimated from Cronbach's alphas, and are considered "good" or "excellent" for all three stakeholders: Candidates ($\alpha = .88$), Faculty Supervisors ($\alpha = .92$), and Mentor Teachers ($\alpha = .93$). Reliability provides an essential estimate of systematic variance, and $1 - \text{reliability}$ provides an operationalization of the percentage of error contained in the scores (Taber, 2018). Based on these reliabilities none of the ProCAD scores contain more than 12% error.

Interrater reliability was measured through two types of agreement: *Absolute* and *Proficiency*. Agreements for each pair of raters (i.e., Teacher Candidate-Faculty Supervisor, Teacher Candidate-Mentor Teacher, Faculty Supervisor-Mentor Teacher) and all three raters (i.e., Teacher Candidate-Faculty Supervisor-Mentor Teacher) were calculated through GoogleTM Sheets for each of the eight ProCAD domains (see Table 4 and Table 5).

Absolute Agreement

Two-Rater, *Absolute* agreements increased from the Middle point (range: 32.97% - 52.75%) to the professional experience's End point (range: 47.44% - 66.67%). In the Middle of the professional experience, the highest *Absolute* Agreements were between candidates and Faculty on the Self-Reflection domain (52.75%). At the End of professional experience, Candidates and Faculty Supervisors still had the highest *Absolute* agreement, but it was on the

Attitudes Towards Students domain (66.67%). At the professional experience's Middle, the lowest *Absolute* agreement was between Faculty Supervisors and Mentor Teachers on Attitudes Towards Students, Towards Schools, and Towards Teaching Methods (32.97%). However, at the End, the lowest *Absolute* agreement was between Teacher Candidates and Mentor Teachers on the Curiosity domain (47.44%). Three-Rater *Absolute* agreement also increased from Middle (range: 15.38% - 28.57%) to the End of the professional experience (range: 30.77% - 46.15%). The highest Three-Rater absolute agreement value was obtained from the Equity dispositions at the Middle (28.57%) and on Open-Mindedness Dispositions at the End of the professional experience (46.15%).

Proficiency Agreement

Across stakeholder combinations, time points, and domains, *Proficiency* agreement was notably higher than *Absolute* agreement. At the Middle of the professional experience Two-Rater *Proficiency* agreements ranged from 70.33% to 90.11%, but it increased by the professional experience's End (range: 87.18% to 98.72%). At the End of the professional experience, the highest proficiency agreements were between Teacher Candidates and Faculty Supervisors as well as Candidates and Mentor Teachers (98.72%) specifically for the domains of Open-Mindedness and Self-Reflection. Three-Rater *Proficiency* agreements increased from the Middle (range: 60.44% to 81.32%) to the professional experience's End point (range: 84.62% to 96.15%). At the Middle, the highest Three-Rater *Proficiency* agreement was on Attitudes Towards Teaching Methods, which had one of the *lowest Absolute*, Three-Rater agreements (15.38%) at the Middle timepoint. At the End of the professional experience, the highest Three-Rater, *Proficiency* agreement was for Open-Mindedness Dispositions which yielded the highest *Absolute* agreement at the End. The lowest Three-Rater, *Proficiency* agreement was obtained on

the Professionalism domain both at the Middle (60.44%) and at the End point (84.62%). In summary, both types of agreement percentages increased from the Middle point to the End point.

Research Question 3: Agreement Across Rater Groups: Intraclass Correlations

Coefficients (ICC)

In addition to Cronbach's Alphas and percentage agreements reported above evaluating response consistency, ICC estimates and their 95% confident intervals were calculated based on a one-way, random-effects model. ICC provided evidence of the magnitude of the relationships based on multiple rater groups. General guidelines for ICC suggest the use of at least three rater groups and 30 members per group (Koo & Li, 2015, p. 158). At the Middle of the professional experience, the ICCs for all eight items representing the domains are considered "poor" based on descriptor conventions for ICCs (see Table 6; Koo & Li, 2015). However, all ICC's for the rater groups increased during the professional experience, from Middle to End, and all are statistically significant at the experience's End. At the End of the professional experience, the ICC for the Open-Mindedness Disposition and the Professional Competencies increased to a "moderate" level of reliability (Koo & Li, 2015). Although all others remained in the "poor" range all of the domain scores are within the 95% Confidence Interval designated as "moderate" (Koo & Li, 2015).

Two-Way, Repeated-Measures ANOVA

In order to test differences in dispositional ratings based on time and stakeholder, a two-way, repeated measures ANOVA was conducted. The within-subjects factor was time (Middle, End) and the between-subjects factor was stakeholder (Teacher Candidate, Faculty Supervisor, and Mentor Teacher) with an outcome variable of dispositional ratings.

Prior to conducting the two-way, repeated-measures ANOVA, various assumptions were checked. Approximate normality was measured through qualitative inspection of histograms where the dispositional ratings were slightly negatively skewed but with similar shapes for both time points, as described above. In fact, skewness and kurtosis were both less than $|1|$ and provide evidence that the distributions are approximately normal distributions. Z-score for the beginning and the End timepoints were saved and frequency tables were created. No scores from the Middle and End times were considered outliers. (Tabachnick & Fidell, 2007).

To test the assumption of equality of error variance, Levene's test was conducted. While significant differences were not detected for the End of professional experience ratings [$F(2, 266) = 2.58, ns$], the error variances were considered significantly different across groups for the Middle of professional experience ratings, $F(2, 266) = 5.10, p = 0.007$. However, given that the distributions contained no outliers and the smallest and largest sample sizes were relatively equal (< 4 to 1) the F Max Ratio of 10 to 1 was used as a test of the equality of variances (Tabachnick & Fidell, 2007). The ratio of the largest to the smallest standard deviation was 1.29 to 1, which falls below the necessary F Max Ratio. Consequently, the variances were not sufficiently different from one another to invalidate the use of parametric statistics .

Research Question 4: Do Ratings Differ Across Time?

Results from the two-way, repeated-measures ANOVA yielded significant main and interaction effects. Main effects of time [$F(1,266) = 57.39, p < 0.001$, partial $\eta^2 = 0.18$] and stakeholder [$F(2, 266) = 12.89, p < 0.001$, partial $\eta^2 = 0.09$] were significant with medium effect sizes; based on the significant F values for the main effects there is evidence that stakeholders' ratings differed across time and as a function of group. However, as indicated by the statistically significant interaction, $F(2, 266) = 15.01, p < 0.001$, partial $\eta^2 = 0.1$,

dispositional ratings depend on both stakeholder and time . See Figure 1 for a depiction of the scores across time by rater. The effect size of this interaction difference is considered “medium” based on eta-squared (Tabachnick & Fidell, 2007). Significant main effects should not be interpreted independently but are best interpreted within the context of the interaction effects. Initially, simple effects analyses were conducted to explore the interaction of time and rater, as described below.

Follow up post-hoc pairwise comparisons with Bonferroni correction were conducted. Candidates’ ($p = 0.048$), Faculty Supervisors’ ($p < 0.001$), and Mentor Teachers’ ($p = 0.016$) dispositional ratings significantly increased from the Middle to the End of the professional experience (see Table 7 for M and SD). However, Faculty Supervisor’s ratings changed the most from the Middle ($M = 23.57$, $SD = 3.47$) to the End ($M = 27.96$, $SD = 3.77$) with an increase of approximately 4.39-points. At the Middle, the average total was slightly below Proficient in the Developing range, but at the End it was within the Proficient range. These descriptors were determined by dividing the average total by the number of domains (see Table 8). Teacher candidates and Mentor Teachers increased by 0.93 points and 1.43 points, respectively.

Research Question 5: Do Stakeholders’ Mean Scores Differ?

Additional, post-hoc, pairwise comparisons were conducted to determine whether rater means differed as a function of the time the observations were taken (i.e., Middle and End). Based on results of a one-way ANOVA, there were significant mean differences in dispositional ratings based on stakeholder, $F(2, 266) = 27.42$, $p < 0.001$ at the Middle time point. More specifically, based on pairwise comparisons, marginal mean differences with a Bonferroni correction, Candidates’ ($M = 26.94$, $SD = 3.26$; $p < 0.001$) and Mentor Teachers’ mean ratings ($M = 26.88$, $SD = 3.47$; $p < 0.001$) were significantly higher than Faculty Supervisors’ ratings ($M =$

23.57, $SD = 4.21$), However, at the End of the professional experience, no stakeholders' ratings were significantly different from one another, $F(2, 266) = 0.31, p = ns$. Therefore, while stakeholders rated candidates' dispositions significantly differently at the Middle, their ratings were not significantly different at the End of the professional experience (See Figure 1).

Chapter IV

Discussion

This study's purpose was to explore further the ProCAD instrument's psychometric properties and utility. Data from 189 Teacher Candidates who completed their professional experience during the 2018-2019 school year were included in these analyses. The psychometric properties were investigated through exploratory factor analysis, reliability analyses, *Absolute* and *Proficiency* percent agreements, and intraclass correlation coefficients. In addition, various other questions related to ProCAD's utility were analyzed via multiple ANOVAs.

Based on exploratory factor analysis, this instrument is best conceptualized as having one general factor, despite having a conceptual or theoretical three-factor structure. Moreover, the highest level of agreement between raters comes from *Proficiency* agreements as opposed to *Absolute* and/or ICC. Lastly, conclusions based on the ANOVAs indicate that the mean stakeholders' ratings increase across time both generally and for each stakeholder and that Teacher Candidates and Mentor Teachers mean ratings differed from the Faculty Supervisors' mean rating at the Middle point but were not significantly different at the End of the professional experience. This study's results, implications, and limitations are discussed below along with suggestions for future research.

Research Question 1: ProCAD Structure

Knowing an instrument's statistically-based structure is necessary in order to make empirically-informed decisions for both statistical analyses and for practical decisions: For example, is it appropriate to use the total ProCAD score as a unitary measure in statistical analyses (e.g., as the dependent variable in an ANOVA context)? Is a candidate's total score an accurate indication of the individuals' dispositions? Does the psychometric integrity of the

ProCAD score justify using Mentor Teachers' and Faculty Supervisors' overall ratings to define candidates' performance? The answer to these questions appears to be a tentative "yes." The results from the study that address these issues and more are discussed below.

Despite the ProCAD instrument's conceptual configuration with three underlying elements (i.e., attitudes, dispositions, and competencies; Laughter et al., 2020), the EFA results yielded a single, underlying structure. Henry and others (2013) also conceptualized their instrument with multiple underlying factors, but empirical evidence provides the best support for a single, general factor structure. In general, rather than using subcomponents of ProCAD, the most defensible conceptualization of dispositional performance is as a total score (Henry et al., 2013; Saltis et al., 2021). However, a caveat is important to consider. Given that the factor analysis relies on group data, the one-factor interpretation may not hold for all, and specifically for those individuals who exhibit significant variability across the three theoretical ProCAD elements: Attitudes, Dispositions, and Competencies. So, Teacher Candidates and their supervisors should continue to evaluate performance according to the three-component structure for individuals when the scores across the three areas are radically different.

How do these results defining characteristics of ProCAD relate to findings from other dispositional instruments? The data appear mixed, though the evidence from some instruments provide support for a multi-factor structural interpretation. Data taken from Choi and other's (2016) Teacher Education Dispositions rating form are less than definitive. The initial fit from an exploratory factor analysis yielded a somewhat similar general one-factor structure to the ProCAD instrument. Further, although based on CFA fit statistics (followed by bifactor analysis), its best fit resulted in a general factor, the data also provided some support for *two* specific dispositional factors (Choi et al., 2016). Other instruments notably departed from the

single factor structure. West and others (2020) created a 26-item measure that best fit with five subscales based on EFA. However, these results may be compromised by the methodology. For example, this study relied on maximum likelihood estimation to fill in missing data without specifying the missingness of the data. Consequently, if a substantial number of cases were missing, as is common in applied educational research, then MLE could have biased the factor analysis (Garson, 2015). So, while some dispositional instruments appear to define multiple factors this configuration is not the exclusive norm for the field. One reason for discrepant findings is probably related to a methodological decision during instrument development. In general, factor structure relies in part on the number of items included—instruments with more items (than eight) tend to yield multiple factors.

Research Question 2: ProCAD Reliability Estimates

ProCAD's overall scores' high internal reliability estimates based on Cronbach's Alphas provided evidence that it can best operationalized as one construct for most Teacher Candidates. As noted above, the internal consistency reliabilities for ProCAD across raters range from .88 – .93, indicating little error variance (i.e., 12% or less). These values provide evidence to support using the overall ProCAD score as a good operationalization of a Teacher Candidates' dispositional performance (Cronbach, 1990).

Similar tools such as the Professional Dispositions Qualities ($\alpha = .85$; Saltis et al, 2021), The Candidate Beliefs SAS ($\alpha = .81$), and the CDPA ($\alpha = .96$; Nweke et al., 2019) all reported Cronbach's alphas similar to the ProCAD instrument (range: $\alpha = .88 - .93$). While other pre-established measures, such as the BATS2 (Lang et al., 2018) report "excellent" reliability without exact alphas, such general reports are minimally informative due to the variability in alpha classification terms across studies and "experts" (Taber, 2018). The similarities to other

instruments in the field, in both measurement strategies and statistical findings, indicate that the ProCAD instrument not only meets the standards accepted in the field but its structure and psychometric integrity as provided by Laughter and others (2020) *and* results from the current study enable test users to better understand and accurately interpret it, although there are caveats.

There is a general understanding in the fields of psychology and education that alphas reflect one critical element of psychometric integrity. Although some other dispositional measures have slightly higher alphas than the ProCAD instrument, that statistic does not necessarily indicate that such instruments are “better” or more “accurate” than the ProCAD instrument. Cronbach’s alpha can be increased through simply adding items to a scale (Cronbach, 1951). If a measurement’s alpha is considered “too small,” then the authors may add more items. This method becomes problematic when items become redundant and, in turn, the measure becomes too lengthy and inefficient (Cronbach, 1951). In these cases, the alpha may be strong, but added questions are not measuring unique dispositions. Instead, the added time and decreased efficiency in completion may discourage individuals from finishing the measure which will result in incomplete data. For example, while the CDPA has an alpha of 0.96, it also has 24 indicators, which makes it less efficient than ProCAD. Insofar as the ProCAD measures all the dispositions Nweke and others (2019) consider important, then it may be considered a more efficient and user-friendly instrument

Given that ProCAD is completed by multiple raters, additional psychometric information can be determined through interrater agreement percentages as well as Intraclass Correlation Coefficients. Evidence of psychometric quality taken at one point in time generalizes to some extent to ratings from other years and raters. Evidence from the data taken from this study provide support for its psychometric integrity based on an additional source of reliability—

interrater agreement. Both *Absolute* and *Proficiency* agreements were measured across stakeholder pairs and three-raters along with across time for each of the eight ProCAD domains. This specific strategy has been used with the ProCAD measure before and was implemented for direct comparison (See Laughter et al., 2020). Notably, the agreements reported in Laughter and others' (2020) article were from the End of the professional experience and only included Two-Rater agreements between Faculty Supervisors and Mentor Teachers in addition to the Three-Rater agreements. In this study, all agreements' ranges were lower at the Middle than at the End of the professional experience. Moreover, these Middle agreement ranges were lower than the End of the professional experience ranges in the previous study (Laughter et al., 2020). At the End of the professional experience, the *Absolute* agreement between supervising faculty and Mentor Teachers in the current study (range: 53.85% - 65.38%) was similar to the previous (range: 54.1% - 64.9%). However, three-rater, *Absolute* agreement was lower in this study (range: 30.77% - 46.15%) compared to the previous evidence (range: 43.7% - 56.8%). When comparing *Proficiency* agreement between Faculty Supervisors and Mentor Teachers at the End of the professional experience, this study (range: 88.46% - 97.44%) was similar to the previous data (range: 90.5% - 93.2%) but with a larger range. Additionally, three-rater proficiency at the End of the professional experience was similar across the three raters in this study (range: 84.62% - 96.15%) and the previous (range: 81.0% - 92.8%), but this study's agreement range was slightly higher. Altogether, not only was the *Proficiency* agreement higher in this study for both two- and three-rater comparisons compared to *Absolute*, but it was also more similar to the previous study's *Proficiency* agreements than the *Absolute* agreements within the previous study.

Research Question 3: Agreement Across Rater Groups: Interrater Correlations

In addition to percentage agreements, the interrelatedness of the stakeholders' ratings was further evaluated through Intraclass Correlation Coefficients, which define shared variance across groups of raters. Although a *one-way, random* intraclass correlation coefficient can account for agreement across groups of different raters, its results are always the smallest out of the ICC options (Landers, 2015). Across domains, the ICC was the highest at the professional experience's End. This study's valid ICCs (range: 0.39 – 0.57) were much lower than those ICCs in Nweke and others' (2019) study (0.80) but are somewhat similar to the Teacher Education Dispositions Rating Form (TEDRF); the TEDRF also had small interrater reliability correlations but high internal consistency across a general factor (Choi et al., 2016).

As mentioned by Portney and Watkins (2015), low coefficients can be accounted for by raters' disagreement; nevertheless, the ICC determines variance across all three rater groups. Because this measure does not indicate the source of the disagreement, only one rater group could disagree with the other two while the other two agree. Based on the additional analyses in this study, the Faculty Supervisors rated candidates lower than Teacher Candidates and Mentor Teachers rated them. Moreover, this difference is most evident at the Middle point, a pattern which may account for the modest ICCs. Given that other statistics (i.e., *Absolute* and *Proficiency Agreement*) and internal consistency reliabilities that are strong, the modest ICCs should be interpreted cautiously.

Research Question 4: Stakeholder Differences

In order to better understand the ProCAD's utility, the stakeholders' ratings were compared and contrasted. At the Middle, Teacher Candidates' and Mentor Teachers' ratings were significantly higher than Faculty Supervisors' ratings, though the mean ratings did not

differ at the End point. Similarly, to results from this study's Middle timepoint, Laughter and others (2020) found that Teacher Candidates rated themselves higher than their instructors in a pilot study with 18 candidates. However, during the full professional experience analyzed in that study with 127 candidates and their supervising mentors, average scores were actually highest for Faculty Supervisors and Mentor Teachers, though there were no statistical tests of mean differences (Laughter et al., 2020). Similar to the pattern in the current study, Saltis and others (2021) found that Mentor Teachers' mean ratings were significantly higher than those of the faculty. Additionally, in a meta-analysis examining ratings from a non-educational, business setting, employees tended to rate themselves higher than their bosses on performance measures, much like the pattern of higher Teacher Candidates' ratings compared to those from Faculty Supervisors (Heidemeier & Moser, 2009). On the other hand, Saltis and others (2021) found that Mentor Teachers' ratings were significantly higher than Teacher Candidates' ratings. Perhaps Faculty Supervisors tend to take on a conceptually similar role to bosses in a business context and provide more rigorous ratings than Mentor Teachers, though the magnitude of the ratings may converge over time and more supervision (as in this study).

Although researching dispositional measures' differences is relatively new, differences among education stakeholders' ratings are not. In the context of lesson plan assessment, Tillema (2009) asked Teacher Candidates, Mentor Teachers, and Faculty Supervisors about their role and the assessment's purpose. The two raters with "authority"—Faculty Supervisors and Mentor Teachers—saw their roles differently. Mentor Teachers saw themselves as "Guides" and "Performance Consultants," whereas Faculty Supervisors saw themselves as "Reflectors" and "Critical Friends" (Tillema, 2009). While they agreed on one purpose of the assessment being "determining progress in development," they judged the importance of other purposes of the

assessment to differ (Tillema, 2009). Whereas Faculty Supervisors saw lesson plan assessments as promoting learning, giving feedback, and determining competency level, Mentor Teachers found these roles to be significantly less important (Tillema, 2009). In addition, across stakeholders, the most commonly noted problem in lesson plan assessment was “Lack of guidelines and grading rules for assessors” (Tillema, 2009). These perceptual differences and common difficulties were all noted within a teacher education context. Although lesson plan assessments are not the *same* as dispositional assessment, incongruent dispositional ratings between stakeholders on the ProCAD instrument may not stem from candidates displaying or possessing different dispositions. Rather, like the raters in the Tillema (2009) study, raters may lack guidance or have differing perceptions of purpose. This difference in role perception may explain why Faculty Supervisors and Mentor Teachers rate candidates so differently: Despite both having a “supervisory” function in the teacher education process, they may perceive themselves as having slightly different roles and goals.

Research Question 5: Changes Over Time

Mean ratings at the end of the professional experience were significantly higher than at the Middle for each of the stakeholders. These results align with much of the literature. Saltis and others (2020) found dispositional rating increases across time points (i.e., program’s beginning, Middle, and End) as did Seay (2017). Unlike those studies, within this study only time-based differences from the professional experience’s Middle to its End were assessed; even so, within this relatively small timespan candidates’ dispositions increased, and ratings converged. These results are not consistent across the literature, however. For example, Immekus (2016) did not find significant dispositional changes over time, which they hypothesized to be due to a misfit between the pre-established instrument and its intended use since the scores lacked variability

even between candidates. Nevertheless, the findings of the current study align with the descriptive statistics found in the previous study of Laughter and others (2020) and is consistent with the perspective of those in the field who view dispositions to not only be important but also malleable (Caroll, 2012; Diez, 2007). While there are those who believe dispositions to be too deeply entrenched to show time-based changes (e.g., McKnight, 2004; Wasicsko et al., 2009), mounting evidence supports the perspective that EPPs can intentionally design and align their coursework with practicum experiences to improve dispositions (e.g., Saltis et al., 2021; Seay, 2017).

In general, ProCAD seems well-suited to measure candidates' dispositional changes over small periods of time and appears sensitive to even small improvements. In addition, it is likely sensitive to more dramatic improvement across longer periods of time (i.e., program beginning to End). Moreover, EPPs should be discouraged from assuming that candidates' dispositions are "set in stone" upon program entrance. Instead, EPPs should intentionally create program coursework and experiences that can help candidates reflect upon their dispositions and can improve them as needed.

Despite differences at the professional experience's Middle, at the year's End, stakeholders' ratings were *not* significantly different. One possibility for this change in ratings is that stakeholders' perceived roles aligned for the final evaluation time, and any issues with misunderstanding or lack of guidance disappeared; however, a more likely explanation for this similarity would be real changes in candidates' dispositions resulting from programmatic design. First, however, the prospective ceiling effect and its possible consequences should be addressed, given that most scores were above 3 on a 4-point scale, as recommended by Vogt (2005); he defined ceiling effect as "what happens when many subjects in a study have scores on a variable

that are at or near the possible upper limit ('ceiling'). Such an effect may cause problems for some types of analysis because it reduces the possible amount of variation" (as cited in Simkovic et al., 2019, p. 3). However, ceiling effects do not appear to be a significant problem for the analyses within this study given the reasonably good estimates of normality (i.e., both skewness and kurtosis estimate were less than $|1|$). In addition, the distributions contained enough variability to yield mean difference analyses that were statistically significant.

The teacher preparation program's goal is for all candidates to reach "Proficiency" (3) with especially meritorious dispositions deserving a rating of "Exemplary" (4; Laughter et al., 2020). Dispositional growth is expected throughout the program's coursework and practical experiences, with "Proficiency" expected by the program's End (Laughter et al., 2020). Moreover, through refinement plans and career counseling, candidates who perform *below* Proficient expectations would have received either intervention and improvement or intervention, no improvement, and removal from the program (Laughter et al., 2020). With these requirements and refinements in mind, similarities between stakeholders scores clustering at the top of the rating scale are not only understandable but also expected.

Limitations and Implications for Future Research and Practice

The current study is characterized by various strengths: proportional representation of the university's different teacher education programs, appropriately large sample sizes, and multiple time points for comparison. While this instrument was specifically developed by and for this university in the southeastern United States, other programs in varying locations with differing student populations have and could inquire about using the instrument. Results provide support for the psychometric integrity and utility of ProCAD. Nonetheless, there are limitations. For example, the participant pool is homogeneous. This study and a relevant previous one that

focused on the ProCAD (Laughter et al., 2020) were conducted at the same state school with the same limited representation of racial and ethnic membership. In addition, the Teacher Candidates have a higher pass rate on qualifying assessment scores than their peers' average throughout the state (Tennessee State Board of Education, 2019). Perhaps these candidates possess unique skills that predispose them to higher qualifying assessment performance and a better aptitude for improving dispositions than their peers at other state institutions. These characteristics limit generalizability. With these limitations in mind, additional analyses of the instruments' psychometric properties and utility should be conducted at different EPPs in order to examine generalizability.

Within this study, the ICCs between stakeholder were relatively weak compared to guidelines in the field (Koo & Li, 2016), indicating limited group-based agreement. However, the *Absolute* agreement ranges were similar to those from a previous study of the ProCAD instrument (Laughter et al., 2020), and the *Proficiency* agreement between all combinations of raters was deemed acceptable. These variations in correlations and agreements paired with the Faculty Supervisors' ratings being lower than the Teacher Candidates' and the Mentor Teachers' ratings at the Middle, indicate that more research into *why* these differences emerge between stakeholders' ratings is needed and particularly exploration of whether raters have a common understanding of goals/roles (Conderman & Walker, 2015; Tillema, 2009).

Within this study, only changes over a short period of time were examined. Unlike other studies in the area of candidate dispositions (e.g., Saltis et al., 2021; Seay, 2017), this study of the ProCAD instrument and the last one by Laughter and others (2020) only considered disposition ratings and changes during the Teacher Candidates' professional year. The ProCAD instrument may be insensitive to longer-term changes. One limitation of obtaining such

longitudinal data is related to matching candidates over longer periods. For example, it is difficult to match candidates over extended time due to various factors (e.g., name changes, misspellings, stakeholders fail to complete form; see Henry et al., 2013) though such an exploration could better inform the nature of Teacher Candidates' dispositional changes and the potential catalysts for improvements (i.e., assignments, courses, practicum experiences).

This study included institutional data and did not include a control group. Future researchers should include a control group if possible, perhaps another cohort of graduate students in a related, or perhaps even an unrelated, field in order to measure changes in self and supervisor ratings. Another field, with no curriculum aimed at improving dispositions, may provide an ideal control group. If the control group shows significant improvements over time, then the dispositional increases could simply be a result of history, maturation, or other threats to validity (Campbell et al., 1963).

An additional area of exploration might provide a more molecular perspective as the psychometric integrity and utility of the ProCAD for Teacher Candidates may differ based on their licensure subject areas. While professional dispositions are important for all teachers to possess, teachers in specific subject areas (e.g., English as an Additional Language, Special Education) may utilize some dispositions more than teachers in other areas. For example, perhaps EAL and SPED teachers need to utilize their knowledge obtained from mastery of dispositions more frequently than other teachers given the diversity of the student population with whom they work. However, given implementation of Multi-Tiered Systems of Support (MTSS) in many classrooms, awareness of students' abilities and disabilities within the same general education classroom is becoming increasingly necessary (Sailor et al., 2021).

Nonetheless, future research into dispositional mastery across program areas may inform the extent to which Teacher Candidates dispositional knowledge varies across subdisciplines.

While the catalyst for creating the psychometrically-sound, ProCAD instrument originated with the requirements from CAEP, that does not mean that the utility of such instruments stops after each accreditation cycle. Various concerns exist within our K-12 education system. Not only is there a teacher shortage but there is also low racial and ethnic diversity within this field (National School Board Association [NSBA], 2019; U.S. Department of Education et al., 2016). By exploring these issues (e.g., teacher shortages and racial underrepresentation), then various uses for dispositions may be explored such as intentionally teaching dispositions focused on teacher retention through bolstering educators' resilience and diversity enhancement . This effort is particularly relevant for reducing K-12 teacher shortages as reported by NSBA (2019). Problematically, nationwide teacher education enrollment was down about 40% according to a recent estimate with teacher attrition occurring at alarming rate (Carver-Thomas & Darling-Hammond, 2019; Collins & Schaaf, 2020; NSBA, 2019). Moreover, teachers must not only be motivated to enter the field but also to stay in it while building resilience and reducing burnout. Teachers who remain in the same position benefit students, specifically, and schools, generally (Carver-Thomas & Darling-Hammond, 2017). Previous recruiting and retention efforts in California and North Carolina show that many teachers prefer education-based careers, but when incentives (e.g., recruitment programs) and supports (e.g., professional training) were removed, teachers also left (Berry & Shields, 2017). So, effective intrinsic supports (e.g., dispositions), as opposed to extrinsic motivators, may be the more effective way to increasing teacher retention and providing students with the added benefits of experienced educators.

In addition to teacher attrition, some Teacher Candidates drop out of preparation programs prior to ever running their own classrooms. Compared to 2010, in 2017, the number of individuals enrolling in and completing teacher education programs decreased (NSBA, 2019). Teacher Candidate enrollment decreased from 719,081 in 2010 to 441,439 in 2017 (NSBA, 2019). Hence, nationwide teacher education enrollment was down about 40% (NSBA, 2019). Moreover, in 2010 there were 232,707 completers, but there were only 159,598 completers in 2017 (30% decrease; NSBA, 2019). Considering the issues surrounding teacher shortages, this low Teacher Candidate retention adds a layer of concern to the already problematic retention levels of school-based teachers. These statistics highlight the importance of motivating pre-service teachers to stay in the field and to build resilience and reduce burnout. Raising consciousness about the value of dispositions may have a beneficial effect on these goals (Kelley, 2004). In the future, researchers should investigate the extent to which ProCAD informs resilience.

In addition to concerns surrounding teacher shortages, teachers' racial makeup does not reflect students' identities. Approximately 54% of students are non-white but only about 15% of teachers identify as non-white (U.S. Department of Education et al., 2016). While teacher racial and ethnic diversity has increased—both in percentage and in number—the ratio of minority teachers to all teachers does not reflect the ratio of minority students to all students (Ingersoll et al., 2017). Additionally, working in diverse schools increases teachers' likelihood of leaving the profession early (Redding & Smith, 2016).

Since most teachers are white but approximately one-half of the student population nationally identifies as Black or Indigenous People of Color (BIPOC), current teachers should work to overcome any implicit biases that might exist. Teachers need to intentionally create

relationships with all students while also purposefully developing dispositions that lead to acceptance of diverse students (Cunningham et al., 2016; Sheng & Han, 2012). Understanding students' cultures and conveying this comprehension allows teachers to connect with and, in turn, to more effectively teach students (Toldson & Lewis, 2013). The highest achieving Black male students (top 20%) at one school identified the ability to trust their teachers as the most important support of their success (Byers, 2018). Additionally, students are less likely to display defiant behavior when their teachers are trustworthy and approach discipline relationally as opposed to authoritatively (Gregory & Ripski, 2008). These examples illustrate again the importance of accurately measuring dispositions in order to see if these are, in fact, connected to desirable outcomes such as feelings of acceptance for diverse students and increases in experience for their teachers.

One means to improving students' experiences and teacher retention could be through cultivating healthy dispositions during candidates' training experiences while also instilling a desire to improve once in the field (Darling-Hammond, 2006; Talbert-Johnson, 2006). In addition, within the increasingly complex and demanding landscape of education, there may be linkages between enhancing motivation and resilience in teachers—both in-service and pre-service—by attending to and improving dispositional characteristics (Darling-Hammond, 2006; Talbert-Johnson, 2006; Phelps & Benson, 2012). However, in order to make empirically-strong inferences about such connections, psychometrically-sound dispositional measures are necessary. Results of this study provide support for the psychometric integrity of the ProCAD, which is an excellent example of a potentially useful and efficient measure of dispositions

Summary and Conclusion

In general, results support the psychometric integrity of the ProCAD, and its utility. Apparently, ProCAD results are best interpreted for most users from a singular factor model. Additionally, *Proficiency* agreement for each of the eight domains was higher than *Actual* agreement across all combinations of stakeholders. Although the internal consistency reliability of ProCAD is strong and addresses its integrity as a solid measure of dispositions, its intraclass Correlation Coefficients between the three rater groups for each of the domains were relatively weak, particularly at the Middle time point. Perhaps the candidates' *proficiency* may be a more defensible measure of their actual dispositions compared to their specific score (1 to 4). The relatively low ICC does not indicate an intrinsically inadequate measure but rater differences for individual Teacher Candidates should be taken into account when ProCAD is used. Lastly, ratings differed not only by time but also by stakeholder. Not only did ProCAD total ratings increase, overall, from Middle to End of the professional experience but these increases were also significant for each of the three stakeholders. While Faculty Supervisors' ratings were significantly lower than Teacher Candidates and Mentor Teachers at the Middle, all three stakeholders' ratings were *not* significantly different at the End of the professional experience.

Overall, results provide evidence to the support ProCAD reliability, and utility. To increase the generalizability of these findings, similar research with the ProCAD instruments is suggested across varying locations with differing candidate demographics, across longer time spans with various intervening activities, and with a control group. Additional research regarding the perceived roles of stakeholders during dispositional assessment along with connections to outcomes in the field (e.g., retention) could also help inform the utility of the ProCAD instrument.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.
- Arnstine, D. (1967). *Philosophy of education: learning and schooling*. New York: Harper & Row.
- Atwater, L. E., & Brett, J. F. (2002). Antecedents and consequences of reactions to developmental 360 feedback. *Journal of Vocational Behavior*, 66, 532-548.
- Bercaw, L. A., Summers, D. G., Colby, S. A., & Payne, M. (2012). Mirror Images: Teacher Candidates' Perceptions of Disposition Development. *Action in teacher education*, 34(5-6), 514-525. doi:10.1080/01626620.2012.729481
- Berry, B. & Shields, P.M. (2017). Solving the teacher shortage: Revisiting the lessons we've learned. *Kappan*, 98(8), 8-18.
- Beverly, C., Santos, K., & Kyger, M. (2006). Developing and Integrating a Professional Disposition Curriculum into a Special Education Teacher Preparation Program. *Teacher education and special education*, 29(1), 26-31. doi:10.1177/088840640602900104
- Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy*, 1(2), 176-216.
- Brandt, J.-O., Bürgener, L., Barth, M., & Redman, A. (2019). Becoming a competent teacher in education for sustainable development: Learning outcomes and processes in teacher education. *International Journal of Sustainability in Higher Education*, 20(4), 630-653. doi:10.1108/IJSHE-10-2018-0183

- Brown, A. & Laughter, J. (Feb. 2020). Professional characteristics, attitudes, and dispositions [Measurement Instrument]. University of Tennessee.
- Byers, J. (2018). The Impact of Teacher Trust on the Achievement Gap Among African American Male Students. ProQuest Dissertations Publishing.
- Council for the Accreditation of Educator Preparation. (2019). 2013 CAEP Standards. Retrieved from caepnet.org
- Council for the Accreditation of Educator Preparation. (2020, November 13). Fall 2020 accreditation council press release. <http://caepnet.org/about/news-room/f2020-national-press-release>
- Council for the Accreditation of Educator Preparation. (2020). 2022 Initial Level Standards. <http://caepnet.org/~media/Files/caep/standards/2022-initial-standards-1-pager-final.pdf?la=en>
- Campbell, D. T., Stanley, J. C., & Gage, N. L. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Carroll, D. (2012). Examining the development of dispositions for ambitious teaching: One teacher candidate's journey. *The New Educator*, 8(1), 38–64. <https://doi.org/10.1080/1547688X.2011.619950>
- Carver-Thomas, D. & Darling-Hammond, L. (2017). Teacher turnover: Why it matters and what we can do about it. Palo Alto, CA: Learning Policy Institute. <https://learningpolicyinstitute.org/product/teacher-turnover>.
- Carver-Thomas, D., & Darling-Hammond, L. (2019). The trouble with teacher turnover: How teacher attrition affects students and schools. *Education Policy Analysis Archives*, 27, 36. doi:10.14507/epaa.27.3699

- Council of Chief State School Officers. (2013). InTASC model core teaching standards and learning progressions for teachers 1.0. https://ccsso.org/sites/default/files/2017-12/2013_INTASC_Learning_Progressions_for_Teachers.pdf
- Council of Chief State School Officers. (2020). Building on Our Momentum. Retrieved from Washington, DC: <https://online.flowpaper.com/753a0706/CCSSOAnnualReport01620SpreadsFinal/#page=1>
- Choi, H., Benson, N., & Shudak, N. (2016). Assessment of Teacher Candidate Dispositions: Evidence of reliability and validity. *Teacher Education Quarterly*, 43(3), 71–89.
- Collins, E., & Schaaf, K. (2020). Teacher retention in Tennessee. Tennessee Department of Education.
- Combs, A. W. (1972). Some basic concepts for teacher education. *Journal of Teacher Education*, 23(3).
- Conderman, G., & Walker, D. A. (2015). Assessing Dispositions in Teacher Preparation Programs: Are Candidates and Faculty Seeing the Same Thing? *The Teacher Educator*, 50(3), 215-231. doi:10.1080/08878730.2015.1010053
- Cronbach, L. (1990). *Essentials of psychological testing*. New York, NY: Harper & Row.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. doi:10.1007/bf02310555.
- Cunningham, W., Nezlek, J., & Banaji, M. (2016). Implicit and Explicit Ethnocentrism: Revisiting the Ideologies of Prejudice. *Personality & Social Psychology Bulletin*, 30(10), 1332–1346. <https://doi.org/10.1177/0146167204264654>

- Damon, W. (2016). Dispositions and Teacher Assessment. *Journal of Teacher Education*, 58(5), 365-369. doi:10.1177/0022487107308732
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, 8, 1.
- Darling-Hammond, L. (2006). Constructing 21st-Century Teacher Education. *Journal of Teacher Education*, 57(3), 300-314. doi:10.1177/0022487105285962
- Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, 23(1), 57-77. doi:10.3102/01623737023001057
- Darling-Hammond, L., Chung, R., & Frelow, F. (2002). Variation in teacher preparation: How well do different pathways prepare teachers to teach? *Journal of Teacher Education*, 53(4), 286-302. doi:10.1177/0022487102053004002
- Diez, M. E. (2007). Looking back and moving forward: Three tensions in the teacher dispositions discourse. *Journal of Teacher Education*, 58(5), 388–396.
- edTPA. (2021). *About edTPA*. Retrieved from https://www.edtpa.com/PageView.aspx?f=GEN_AboutEdTPA.html
- Fleener, C., & Dahm, P. F. (2007). Elementary teacher attrition: A comparison of the effects of professional development schools and traditional campus-based programs. *Teacher Education and Practice*, 20(263-283).
- Flowers, C. (2006). Confirmatory factor analysis of scores on the clinical experience rubric: A measure of dispositions for preservice teachers. *Educational and Psychological Measurement*, 66(3), 478–488. <https://doi.org/10.1177/0013164405282458>

- Frederiksen, H., Cooner, D., & Stevenson, C. (2012). Assessing teacher dispositions in pre-service teachers. *Journal of College Teaching & Learning (TLC)*, 9(1), 39-52.
- Freeman, L. (2007). Overview of Dispositions in Teacher Education. In J. s & M. E. Diez (Eds.), *Dispositions in Teacher Education* (pp. 3-30). Charlotte, North Carolina: Information Age Publishing.
- Fonseca-Chacana, J. (2019). Making teacher dispositions explicit: A participatory approach. *Teaching and Teacher Education*, 77, 266–276.
<https://doi.org/10.1016/j.tate.2018.10.018>
- Garson, G.D. (2015). *Statistical Associates Blue Book Series: Missing Values & Data Imputation*. Statistical Associates Publishing.
- Ginsberg, R. & Whaley, D. (2003). Admission and retention policies in teacher preparation programs: Legal and practical issues. *The Teacher Educator*, 38(3)169-189, DOI: 10.1080/08878730309555316
- Gregory, A. & Ripski, M. B. (2008). Adolescent trust in teachers: Implications for behavior in the high school classroom, *School Psychology Review*, 37(3) 337-353.
- Guzzo, V. (2018, May 16). Watermark announces partnership with the EDA, LLC to support formal dispositions assessment. Watermark.
<https://www.watermarkinsights.com/press/eda/>
- Hallam, M. K. (2009). Another piece of the language puzzle: Why teacher dispositions are a crucial aspect of student success. *The Language Educator*.
- Harrison, J., Smithey, G., McAfee, H., & Weiner, C. (2006). Assessing candidate disposition for admission into teacher education: Can just anyone teach? *Action in teacher education*, 27(4), 72-80. doi:10.1080/01626620.2006.10463403

- Heidemeier, H., & Moser, K. (2009). Self-Other Agreement in Job Performance Ratings: A Meta-Analytic Test of a Process Model. *Journal of applied psychology*, 94(2), 353-370. doi:10.1037/0021-9010.94.2.353
- Henry, G., Campbell, S., Thompson, C., Patriarca, L., Luterbach, K., Lys, D., & Covington, V. (2013). The predictive validity of measures of teacher candidate programs and performance: Toward an evidence-based approach to teacher preparation. *Journal of Teacher Education*, 64(5), 439–453. <https://doi.org/10.1177/0022487113496431>
- Immekus, J. (2016). The use of surveys in teacher education programs to meet accreditation standards: preservice teachers' culturally responsive beliefs and practices. *Research & Practice in Assessment*, 11(1), 18–28.
- Ingersoll, R., Merrill, L., Owens, C., & Zuckerberg, A. (2017). A quarter century of changes in the elementary and secondary teaching force: From 1987 to 2012. Retrieved from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2017092>
- Interstate New Teachers Assessment and Support Consortium (1992). Model Standards for Beginning Teacher Licensing, Assessment and Development: A Resource for State Dialogue. Retrieved from <https://files.eric.ed.gov/fulltext/ED369767.pdf>
- Katz, L. G., & Raths, J. D. (1985). Dispositions as goals for teacher education. *Teaching and teacher education*, 1(4), 301-307. doi:10.1016/0742-051X(85)90018-6
- Kee, A. N. (2012). Feelings of preparedness among alternatively certified teachers: What is the role of program features? *Journal of Teacher Education*, 63(1), 23-38. doi:10.1177/0022487111421933
- Kelley, L. M. (2004). Why induction matters. *Journal of Teacher Education*, 55(5), 438-448. doi:10.1177/0022487104269653

- Koo, T., & Li, M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163.
<https://doi.org/10.1016/j.jcm.2016.02.012>
- Landers, R. (2015). Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS. *The Winnower*, 2. DOI:10.15200/winn.143518.81744
- Lang, W.S., Moore, L., Wilkerson, J.R., Parfitt, C.M., Greene, J., Kratt, D., Martelli, D., LaPaglia, K., Johnston, V., Gilbert, S., Zhang, J., Fields, L. (2018). Beliefs about Teachings (BATS2) - Construction and validation of an instrument based on InTASC critical dispositions. *International Journal of Learning, Teaching and Educational Research*, 17(8).
- Latham, N. I., & Vogt, W. P. (2007). Do professional development schools reduce teacher attrition? Evidence from a longitudinal study of 1,000 graduates. *Journal of Teacher Education*, 58(2), 153-167. doi:10.1177/0022487106297840
- Laughter, J. C., Brown, A. A., & Neu, L. W. (2020). ProCAD: Looking at a fuller picture of the teacher candidate. *Journal of Teacher Education*, 0022487120972635.
- Leibbrand, J. (2010). NCATE and TEAC Form New Accrediting Body: The Council for the Accreditation of Educator Preparation (CAEP) [Press release].
- Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
<https://doi.org/10.1080/01621459.1988.10478722>
- Maurer, T. J., Mitchell, D. R. D., & Barbeite, F. G. (2002). Predictors of attitudes toward a 360-degree feedback system and involvement in post-feedback management development

- activity. *Journal of Occupational and Organizational Psychology*, 75(1), 87-107.
- doi:10.1348/096317902167667
- McKnight, D. (2004). An inquiry of NCATE's move into virtue ethics by way of dispositions (Is this what Aristotle meant?). *Educational Studies*, 35, 212-230.
- Minnesota Higher Education Coordinating Board, S. P. (1986). Minnesota's Vision for Teacher Education: Stronger Standards, New Partnerships. Report of the Task Force on Teacher Education for Minnesota's Future.
- Morrow, J. (2018, February 18). *Mixed ANOVA output* [PowerPoint Slides]. UTK Canvas.
- <https://utk.instructure.com/courses/45511>
- Morshead, R. W. (1968). Philosophy of education: Learning and schooling. [Review of Philosophy of education: Learning and schooling by D. Arnstine]. *Studies in Philosophy and Education*, 6, 91–105. doi:<https://doi.org/10.1007/BF00367801>
- National School Board Association. (2019, May 20). Fixing the holes in the teacher pipeline: Three ways for school leaders to act. <https://nsba.org/Perspectives/2019/Fixing-Teacher-Pipeline-Holes>
- National Council for Accreditation of Teacher Education. (2008). Professional Standards for the Accreditation of Teacher Preparation Institutions. Washington, D.C.
- Nweke, W. C., Perkins, T. P., & Afolabi, C. Y. (2019). Reliability analysis of complementary assessment tools for measuring teacher candidate dispositions. *Georgia Educational Researcher*, 16(2). Retrieved from <https://login.proxy.lib.utk.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ1228225&site=ehost-live&scope=site>
- Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin*, 88(639), 60-79.

- Pett, M., Lackey, N., & Sullivan, J. (2003). Making sense of factor analysis the use of factor analysis for instrument development in health care research . SAGE.
- Phelps, P. H., & Benson, T. R. (2012). Teachers with a passion for the profession. *Action in Teacher Education*, 34(1), 65-76. doi:10.1080/01626620.2012.642289
- Portney, L. G., & Watkins, M. P. (2008). *Foundations of clinical research : applications to practice* (3rd ed.). Pearson/Prentice Hall.
- Redding, C. & Smith, T.M. (2016). Easy in, easy out: Are alternatively certified teachers turning over at increased rates? *American Educational Research Journal*, 53(4). 1086–1125
- Rockoff, J. E., & Speroni, C. (2011). Subjective and objective evaluations of teacher effectiveness: Evidence from New York City. *Labour Economics*, 18, 687-696.
- Sailor, W., Skrtic, T. M., Cohn, M., & Olmstead, C. (2021). Preparing teacher educators for statewide scale-up of multi-tiered system of support (MTSS). *Teacher Education and Special Education*, 44(1), 24–41. <https://doi.org/10.1177/0888406420938035>
- Saltis, M.N., Giancaterino, B. & Pierce, C. (2021) Professional Dispositions of Teacher Candidates: Measuring dispositions at a large teacher preparation university to meet national standards, *The Teacher Educator*, 56(2), 117-131, DOI: 10.1080/08878730.2020.1817217
- Schulte, L.; Edick, N.; Edwards, S.; and Mackiel, D. (2005) The Development and validation of the Teacher Dispositions Index. *Essays in Education*, 12. <https://openriver.winona.edu/eie/vol12/iss1/7>
- Schussler, D. L., Stooksberry, L. M., & Bercaw, L. A. (2010). understanding teacher candidate dispositions: Reflecting to build self-awareness. *Journal of Teacher Education*, 61(4), 350-363. doi:10.1177/0022487110371377

- Seay, D. (2017). An examination of the relationship between a candidate's disposition assessments from admission to the program to completion of the program. (Publication No. 10261931) [Doctoral dissertation, Northcentral University]. ProQuest Dissertations Publishing.
- Sheng, F., & Han, S. (2012). Manipulations of cognitive strategies and intergroup relationships reduce the racial bias in empathic neural responses. *NeuroImage (Orlando, Fla.)*, 61(4), 786–797. <https://doi.org/10.1016/j.neuroimage.2012.04.028>
- Shepard, L. A. (2016). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4-14. doi:10.3102/0013189x029007004
- Simkovic, M., Trauble, B., & Hutson, A. (2019). Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PloS One*, 14(8), e0220889–e0220889. <https://doi.org/10.1371/journal.pone.0220889>
- Siwatu, K. (2007). Preservice teachers' culturally responsive teaching self-efficacy and outcome expectancy beliefs. *Teaching and Teacher Education*, 23(7), 1086–1101. <https://doi.org/10.1016/j.tate.2006.07.011>
- Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics (5th ed.). *Allyn & Bacon/Pearson Education*.
- Taber, K.S. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education* 48, 1273–1296 <https://doi.org/10.1007/s11165-016-9602-2>
- Talbert-Johnson, C. (2006). Preparing highly qualified teacher candidates for urban schools: The importance of dispositions. *Education and Urban Society*, 39(1), 147-160.

- Tennessee State Board of Education. (2019). 2019 Report card: University of Tennessee – Knoxville. <https://www.tn.gov/content/dam/tn/stateboardofeducation/documents/2019-educator-preparation-report-card/2019-provider-reports/UTK%202019.pdf>
- Tillema, H. H. (2009). Assessment for learning to teach: Appraisal of practice teaching lessons by mentors, supervisors, and student teachers. *Journal of Teacher Education*, 60(2), 155-167. doi:10.1177/0022487108330551
- Toldson, I., & Lewis, C. W. (Eds.). (2013). Black male teachers : Diversifying the united states' teacher workforce. Retrieved from <https://ebookcentral-proquest-com.proxy.lib.utk.edu>
- U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. (2016). The state of racial diversity in the educator workforce. Washington, D.C.
- Wasicsko, M. M., Wirtz, P., & Resor, C. (2009). Using dispositions in the teacher admission process. *SRATE Journal*, 18(2), 19–26.
- Watermark. (2021, May 4). Educator Disposition Solutions from Watermark and the EDA. Watermark Insights. <https://www.watermarkinsights.com/educator-disposition-assessment/>
- West, C., Baker, A., Ehrich, J. F., Woodcock, S., Bokosmaty, S., Howard, S. J., & Eady, M. J. (2018). Teacher Disposition Scale (TDS): Construction and psychometric validation. *Journal of Further and Higher Education*, 44(2), 185-200. doi:10.1080/0309877x.2018.1527022
- Whitsett, G., Roberson, T., Julian, K., & Beckham, L. (2007). First year teachers' reported levels of functioning on selected professional dispositions. *Education (Chula Vista)*, 128(1), 95.

Young, A., & Wilkins, E. (2008). Dispositional statements on student teacher evaluation instruments: commonalities across institutions. *The Teacher educator*, 43(3), 198-215.
doi:10.1080/08878730802055115

Appendices

Appendix A

Table 1

Means and Standard Deviations for each ProCAD across the three stakeholders and the two time points.

		Middle							End							Mean Diff (%)
		Rating Frequency							Rating Frequency							
		N	1	2	3	4	Mean	SD	N	1	2	3	4	Mean	SD	
Attitudes Towards Teaching Methods	Candidate	152	--	4	74	74	3.46	0.55	122	--	2	47	73	3.58	0.53	3%
	Faculty Supervisor	142	--	14	94	34	3.14	0.57	119	--	5	47	67	3.52	0.58	12%
	Teacher Mentor	102	1	7	35	59	3.49	0.67	88	--	4	27	57	3.60	0.58	3%
Attitudes Towards Students	Candidate	152	--	6	75	71	3.43	0.57	122	--	5	47	70	3.53	0.58	3%
	Faculty Supervisor	142	--	21	89	32	3.08	0.61	119	--	4	50	65	3.51	0.57	14%
	Teacher Mentor	102	1	10	37	54	3.41	0.71	88	--	6	21	61	3.62	0.61	6%
Attitudes Towards Schools	Candidate	152	--	9	70	73	3.42	0.60	122	--	4	65	53	3.40	0.56	-1%
	Faculty Supervisor	142	1	10	101	30	3.13	0.54	119	--	6	63	50	3.37	0.58	8%
	Teacher Mentor	102	1	10	39	52	3.39	0.71	88	--	9	22	57	3.55	0.68	5%
Open-Mindedness Dispositions	Candidate	152	--	2	75	75	3.48	0.53	122	--	1	58	63	3.51	0.52	1%
	Faculty Supervisor	142	--	22	91	29	3.05	0.60	119	--	3	51	65	3.52	0.55	15%
	Teacher Mentor	102	1	10	41	50	3.37	0.70	88	--	5	26	57	3.59	0.60	7%
Self-Reflection Dispositions	Candidate	152	--	1	63	88	3.57	0.51	122	--	4	44	74	3.57	0.56	0%
	Faculty Supervisor	142	1	12	96	33	3.13	0.57	119	--	5	43	71	3.55	0.58	13%
	Teacher Mentor	102	3	12	37	50	3.31	0.80	88	--	4	30	54	3.57	0.58	8%
Curiosity Dispositions	Candidate	152	--	8	84	60	3.34	0.58	122	--	4	53	65	3.50	0.56	5%
	Faculty Supervisor	142	--	26	92	24	2.99	0.60	119	--	6	49	64	3.49	0.60	17%
	Teacher Mentor	102	1	16	43	42	3.24	0.75	88	--	10	30	48	3.43	0.69	6%
Equity Dispositions	Candidate	152	--	14	89	49	3.23	0.60	122	--	2	76	44	3.34	0.51	3%
	Faculty Supervisor	142	--	29	90	23	2.96	0.61	119	--	6	70	43	3.31	0.56	12%
	Teacher Mentor	102	1	9	54	38	3.26	0.66	88	--	6	32	50	3.50	0.63	7%
Professional Competencies	Candidate	152	--	15	79	58	3.28	0.64	122	--	6	62	54	3.39	0.58	3%
	Faculty Supervisor	142	1	29	77	35	3.03	0.69	119	--	12	38	69	3.48	0.68	15%
	Teacher Mentor	102	1	20	39	42	3.20	0.78	88	--	9	30	49	3.45	0.68	8%

Note: Mean Difference is based on the percentage of positive change in ratings from the Middle to the End of the professional experience. Negative percentages indicate a rating decrease.

Table 2

Frequency Counts of Program Areas in the 2018-2019 professional experience cohort with ProCAD ratings.

Program Area	Count	Percentage
Agriculture	10	5.3%
Art Education	6	3.2%
Deaf Education	10	5.3%
Elementary Education	34	18.0%
English as a Second Language	3	1.6%
Middle Grades Math Education	2	1.1%
Music Education	28	14.8%
PreK-3 Early Childhood	19	10.1%
Secondary English Education and ESL	13	6.9%
Secondary Math Education	2	1.1%
Secondary Science Education	8	4.2%
Secondary Social Science Education	8	4.2%
World Language	1	0.5%
Special Education	26	13.8%
VolsTeach Math	6	3.2%
VolsTeach Science	13	6.9%

Table 3

Loadings of the Three Stakeholder's Scores onto Single, General Factor through EFA with Promax Rotation.

ProCAD Domain	Candidate	Faculty Supervisor	Teacher Mentor
Attitudes Towards Teaching Methods	0.72**	0.76**	0.83**
Attitudes Towards Students	0.70*	0.78**	0.80**
Attitudes Towards Schools	0.67*	0.75**	0.85**
Open-Mindedness Dispositions	0.68*	0.81**	0.81**
Self-Reflection Dispositions	0.64*	0.70**	0.74**
Curiosity Dispositions	0.77**	0.75**	0.88**
Equity Dispositions	0.70*	0.77**	0.69*
Professional Competencies	0.65*	0.87**	0.68*
* "Very Good" loading > 0.55 (Comrey & Lee, 1992 as cited in Tabachnick & Fidell, 2007)			
** "Excellent" loading > 0.71 (Comrey & Lee, 1992 as cited in Tabachnick & Fidell, 2007)			

Table 4*Middle of Professional Experience ProCAD Score Agreements*

ProCAD Domain	Absolute Agreement				Proficiency Agreement			
	Candidate/ Supervisor	Supervisor/ Teacher Mentor	Candidate/ Teacher Mentor	Three- rater	Candidate/ Supervisor	Supervisor/ Teacher Mentor	Candidate/ Teacher Mentor	Three-rater
Attitudes Towards Teaching Methods	49.45%	32.97%	40.66%	15.38%	87.91%	84.62%	90.11%	81.32%
Attitudes Towards Students	42.86%	32.97%	45.05%	17.58%	79.12%	81.32%	89.01%	74.73%
Attitudes Towards Schools	42.86%	32.97%	47.25%	15.38%	82.42%	86.81%	86.81%	78.02%
Open-Mindedness Dispositions	47.25%	38.46%	43.96%	17.58%	87.91%	80.22%	90.11%	79.12%
Self-Reflection Dispositions	52.75%	46.15%	51.65%	27.47%	90.11%	84.62%	85.71%	80.22%
Curiosity Dispositions	50.55%	46.15%	42.86%	23.08%	81.32%	81.32%	80.22%	71.43%
Equity Dispositions	47.25%	47.25%	47.25%	28.57%	73.63%	82.42%	82.42%	69.23%
Professional Competencies	45.05%	35.16%	41.76%	19.78%	74.73%	70.33%	75.82%	60.44%

Table 5*End of Professional Experience ProCAD Score Agreements*

ProCAD Domain	Absolute Agreement				Proficiency Agreement			
	Candidate/ Supervisor	Supervisor/ Teacher Mentor	Candidate/ Teacher Mentor	Three- Rater	Candidate/ Supervisor	Supervisor/ Teacher Mentor	Candidate/ Teacher Mentor	Three-Rater
Attitudes Towards Teaching Methods	56.41%	64.10%	51.28%	37.18%	94.87%	96.15%	93.59%	92.31%
Attitudes Towards Students	66.67%	56.41%	51.28%	38.46%	94.87%	91.03%	93.59%	89.74%
Attitudes Towards Schools	53.85%	53.85%	50.00%	30.77%	92.31%	91.03%	93.59%	88.46%
Open-Mindedness Dispositions	64.10%	65.38%	60.26%	46.15%	98.72%	97.44%	96.15%	96.15%
Self-Reflection Dispositions	56.41%	58.97%	62.82%	39.74%	93.59%	94.87%	98.72%	93.59%
Curiosity Dispositions	61.54%	55.13%	47.44%	33.33%	96.15%	88.46%	87.18%	85.90%
Equity Dispositions	65.38%	55.13%	51.28%	38.46%	94.87%	91.03%	93.59%	89.74%
Professional Competencies	55.13%	58.97%	51.28%	34.62%	89.74%	89.74%	89.74%	84.62%

Table 6*ProCAD Score Intraclass Correlation Coefficients*

ProCAD Domain	Middle*					End**				
	F Test with True Value 0					F Test with True Value 0				
	Intraclass Correlation	95% Confidence Interval	Value	df1/df2	Sig	Intraclass Correlation	95% Confidence Interval	Value	df1/df2	Sig
Attitudes Towards Teaching Methods	-0.10	-0.56/0.24	0.91	90/182	0.70	0.39	0.11/0.59	1.64	77/156	0.005
Attitudes Towards Students	-0.08	-0.52/0.26	0.93	90/182	0.65	0.43	0.17/0.62	1.75	77/156	0.002
Attitudes Towards Schools	0.003	-0.41/0.31	1.004	90/182	0.48	0.46	0.22/0.64	1.86	77/156	<0.001
Open-Mindedness Dispositions	-0.08	-0.54/0.25	0.92	90/182	0.66	0.55	0.35/0.70	2.23	77/156	<0.001
Self-Reflection Dispositions	0.38	0.12/0.57	1.61	90/182	0.003	0.46	0.21/0.64	1.84	77/156	<0.001
Curiosity Dispositions	0.21	-0.12/0.45	1.26	90/182	0.10	0.43	0.17/0.62	1.75	77/156	0.002
Equity Dispositions	-0.23	-0.74/0.15	0.82	90/182	0.86	0.49	0.25/0.66	1.94	77/156	<0.001
Professional Competencies	0.05	-0.35/0.34	1.05	90/182	0.38	0.57	0.37/0.71	2.31	77/156	<0.001
Total Score										

*For all of the Middle ICCs, df1 is 90 and df2 is 182.

** For all the End ICCs, df1 is 77 and df2 is 156.

Table 7

Means and standard deviations for stakeholders' dispositional scores based on time point.

Stakeholder	Middle		End	
	Mean (<i>M</i>)	Standard	Mean (<i>M</i>)	Standard
	Deviation (<i>SD</i>)		Deviation (<i>SD</i>)	
Candidate	26.94	3.26	27.87	3.18
Faculty Supervisor	23.57	3.47	27.96	3.77
Teacher Mentor	26.88	4.21	28.31	3.92

Table 8*Ranges of Total Averages and Domain Averages with ProCAD Descriptors*

	Total	Single Domain
	Range	Range
Unsatisfactory	8 – 15.9	1 – 1.9
Developing	16 - 23.9	2 – 2.9
Proficient	24 – 31.9	3 – 3.9
Exemplary	32	4

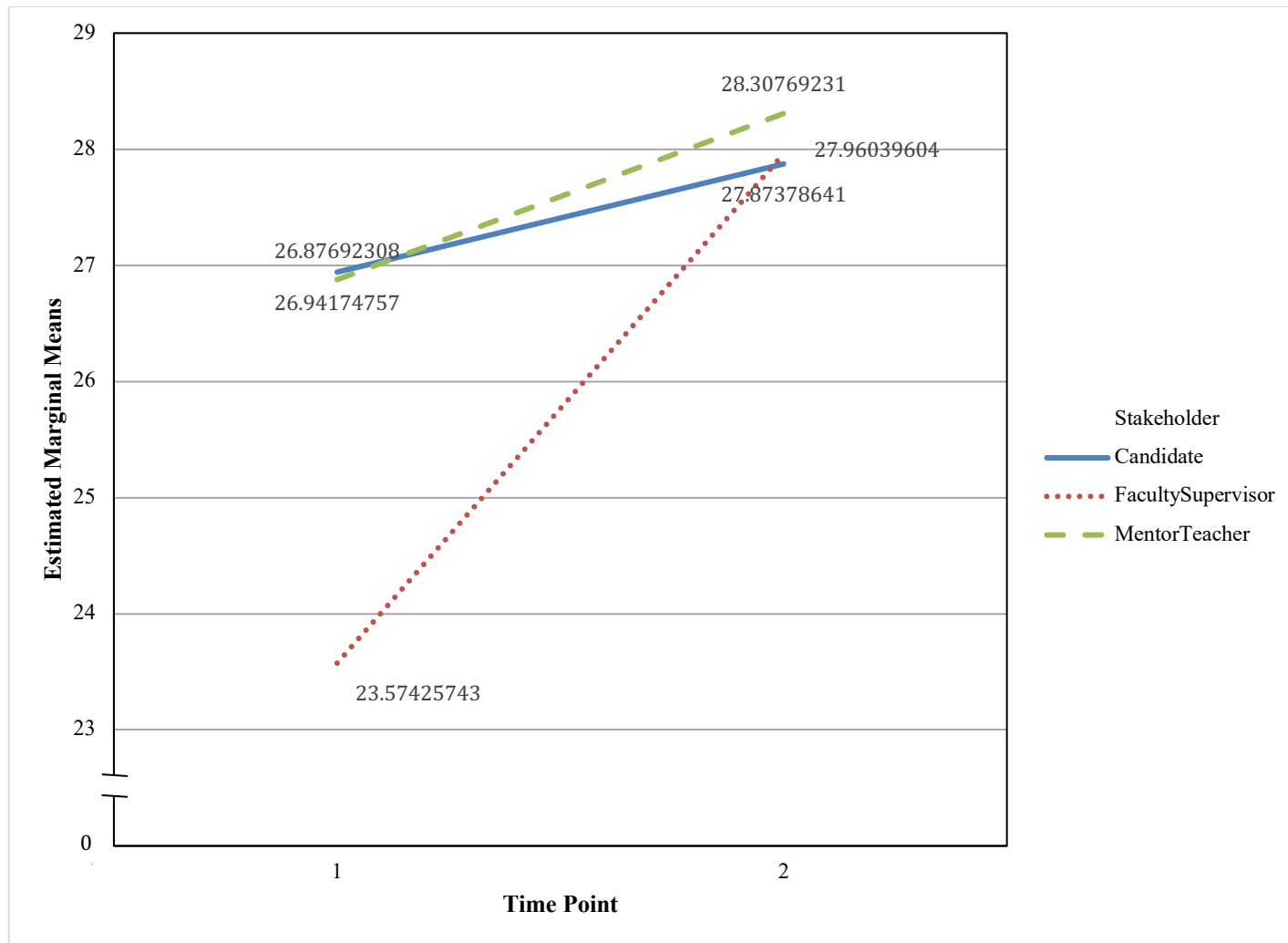


Figure 1

Interaction Effect of Time and Stakeholder on Dispositional Ratings

Appendix B

InTASC Critical Dispositions Sub-Standards

Standard #1: Learner Development

The teacher understands how learners grow and develop, recognizing that patterns of learning and development vary individually within and across the cognitive, linguistic, social, emotional, and physical areas, and designs and implements developmentally appropriate and challenging learning experiences.

- 1(h) The teacher respects learners' differing strengths and needs and is committed to using this information to further each learner's development.
- 1(i) The teacher is committed to using learners' strengths as a basis for growth, and their misconceptions as opportunities for learning.
- 1(j) The teacher takes responsibility for promoting learners' growth and development.
- 1(k) The teacher values the input and contributions of families, colleagues, and other professionals in understanding and supporting each learner's development.

Standard #2: Learning Differences

The teacher uses understanding of individual differences and diverse cultures and communities to ensure inclusive learning environments that enable each learner to meet high standards.

- 2(l) The teacher believes that all learners can achieve at high levels and persists in helping each learner reach his/her full potential.
- 2(m) The teacher respects learners as individuals with differing personal and family backgrounds and various skills, abilities, perspectives, talents, and interests.
- 2(n) The teacher makes learners feel valued and helps them learn to value each other.
- 2(o) The teacher values diverse languages and dialects and seeks to integrate them into his/her instructional practice to engage students in learning.

Standard #3: Learning Environments

The teacher works with others to create environments that support individual and collaborative learning, and that encourage positive social interaction, active engagement in learning, and self motivation.

- 3(n) The teacher is committed to working with learners, colleagues, families, and communities to establish positive and supportive learning environments.
- 3(o) The teacher values the role of learners in promoting each other's learning and recognizes the importance of peer relationships in establishing a climate of learning.
- 3(p) The teacher is committed to supporting learners as they participate in decision making, engage in exploration and invention, work collaboratively and independently, and engage in purposeful learning.
- 3(q) The teacher seeks to foster respectful communication among all members of the learning community.
- 3(r) The teacher is a thoughtful and responsive listener and observer.

Standard #4: Content Knowledge

The teacher understands the central concepts, tools of inquiry, and structures of the discipline(s) he or she teaches and creates learning experiences that make these aspects of the discipline accessible and meaningful for learners to assure mastery of the content.

- 4(o) The teacher realizes that content knowledge is not a fixed body of facts but is complex, culturally situated, and ever evolving. S/he keeps abreast of new ideas and understandings in the field.
- 4(p) The teacher appreciates multiple perspectives within the discipline and facilitates learners' critical analysis of these perspectives.
- 4(q) The teacher recognizes the potential of bias in his/her representation of the discipline and seeks to appropriately address problems of bias.
- 4(r) The teacher is committed to work toward each learner's mastery of disciplinary content and skills.

Standard #5: Application of Content

The teacher understands how to connect concepts and use differing perspectives to engage learners in critical thinking, creativity, and collaborative problem solving related to authentic local and global issues.

- 5(q) The teacher is constantly exploring how to use disciplinary knowledge as a lens to address local and global issues.
- 5(r) The teacher values knowledge outside his/her own content area and how such knowledge enhances student learning.
- 5(s) The teacher values flexible learning environments that encourage learner exploration, discovery, and expression across content areas.

Standard #6: Assessment

The teacher understands and uses multiple methods of assessment to engage learners in their own growth, to monitor learner progress, and to guide the teacher's and learner's decision making.

- 6(q) The teacher is committed to engaging learners actively in assessment processes and to developing each learner's capacity to review and communicate about their own progress and learning.
- 6(r) The teacher takes responsibility for aligning instruction and assessment with learning goals.
- 6(s) The teacher is committed to providing timely and effective descriptive feedback to learners on their progress.
- 6(t) The teacher is committed to using multiple types of assessment processes to support, verify, and document learning.
- 6(u) The teacher is committed to making accommodations in assessments and testing conditions, especially for learners with disabilities and language learning needs.
- 6(v) The teacher is committed to the ethical use of various assessments and assessment data to identify learner strengths and needs to promote learner growth.

Standard #7: Planning for Instruction

The teacher plans instruction that supports every student in meeting rigorous learning goals by drawing upon knowledge of content areas, curriculum, cross-disciplinary skills, and pedagogy, as well as knowledge of learners and the community context.

- 7(n) The teacher respects learners' diverse strengths and needs and is committed to using this information to plan effective instruction.
- 7(o) The teacher values planning as a collegial activity that takes into consideration the input of learners, colleagues, families, and the larger community.
- 7(p) The teacher takes professional responsibility to use short- and long-term planning as a means of assuring student learning.
- 7(q) The teacher believes that plans must always be open to adjustment and revision based on learner needs and changing circumstances.

Standard #8: Instructional Strategies

The teacher understands and uses a variety of instructional strategies to encourage learners to develop deep understanding of content areas and their connections, and to build skills to apply knowledge in meaningful ways.

- 8(p) The teacher is committed to deepening awareness and understanding the strengths and needs of diverse learners when planning and adjusting instruction.
- 8(q) The teacher values the variety of ways people communicate and encourages learners to develop and use multiple forms of communication.
- 8(r) The teacher is committed to exploring how the use of new and emerging technologies can support and promote student learning.
- 8(s) The teacher values flexibility and reciprocity in the teaching process as necessary for adapting instruction to learner responses, ideas, and needs.

Standard #9: Professional Learning and Ethical Practice

The teacher engages in ongoing professional learning and uses evidence to continually evaluate his/her practice, particularly the effects of his/her choices and actions on others (learners, families, other professionals, and the community), and adapts practice to meet the needs of each learner.

- 9(l) The teacher takes responsibility for student learning and uses ongoing analysis and reflection to improve planning and practice.
- 9(m) The teacher is committed to deepening understanding of his/her own frames of reference (e.g., culture, gender, language, abilities, ways of knowing), the potential biases in these frames, and their impact on expectations for and relationships with learners and their families.
- 9(n) The teacher sees him/herself as a learner, continuously seeking opportunities to draw upon current education policy and research as sources of analysis and reflection to improve practice.
- 9(o) The teacher understands the expectations of the profession including codes of ethics, professional standards of practice, and relevant law and policy.

Standard #10: Leadership and Collaboration

The teacher seeks appropriate leadership roles and opportunities to take responsibility for student learning, to collaborate with learners, families, colleagues, other school professionals, and community members to ensure learner growth, and to advance the profession.

- 10(p) The teacher actively shares responsibility for shaping and supporting the mission of his/her school as one of advocacy for learners and accountability for their success.

- 10(q) The teacher respects families' beliefs, norms, and expectations and seeks to work collaboratively with learners and families in setting and meeting challenging goals.
- 10(r) The teacher takes initiative to grow and develop with colleagues through interactions that enhance practice and support student learning.
- 10(s) The teacher takes responsibility for contributing to and advancing the profession.
- 10(t) The teacher embraces the challenge of continuous improvement and change.

(Council of Chief State School Officers, 2013)

Appendix C

Professional Attitudes

A preservice teacher embodies *attitudes* toward a variety of topics and people: attitudes toward teaching methods, attitudes toward students, attitudes toward schools. These attitudes form the internal paradigm through which a preservice teacher interprets the world and the classroom.

Attitudes toward Teaching Methods

Content and assessment change alongside changes in the broader society and a teacher must understand and address those changes through effective planning, instruction, and assessment. Lesson plans and methods should not be understood as guidelines written in stone but a reflection on possibilities that can be adapted to new and diverse situations.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
1(k) The teacher values the input and contributions of families, colleagues, and other professionals in understanding and supporting each learner's development. 2(o) The teacher values diverse languages and dialects and seeks to integrate them into his/her instructional practice to engage students in learning. 5(r) The teacher values knowledge outside his/her own content area and how such knowledge	The teacher expresses a value for only their own knowledge and input in the classroom and is inflexible with their learning environment. (Note: rating of Unsatisfactory may trigger the creation of a deficiency report.)	The teacher expresses a value for the input of multiple stakeholders in the classroom and works to develop an open, safe, and flexible classroom environment.	The teacher expresses a value for the input of multiple stakeholders and generally develops an open, safe, and flexible classroom environment.	The teacher consistently expresses a value for the classroom as a collaborative effort and regularly develops an open, safe, and flexible classroom environment.

enhances student learning. 5(s) The teacher values flexible learning environments that encourage learner exploration, discovery, and expression across content areas. 7(q) The teacher believes that plans must always be open to adjustment and revision based on learner needs and changing circumstances. 8(s) The teacher values flexibility and reciprocity in the teaching process as necessary for adapting instruction to learner responses, ideas, and needs.				
--	--	--	--	--

Attitudes toward Students

Diversity exists across students and teachers and within students and teachers. Effective teaching values that diversity and includes it in all areas of planning, instruction, and assessment.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
2(l) The teacher believes that all learners can achieve at high levels and persists in helping each learner reach his/her full potential. 3(o) The teacher values the role of learners in promoting each other's learning and recognizes the importance of peer relationships in establishing a climate of learning. 8(q) The teacher values the variety of ways people communicate and encourages learners to develop and use multiple forms of communication.	The teacher demonstrates a belief that only some students have the capacity for high academic achievement. (Note: rating of Unsatisfactory may trigger the creation of a deficiency report.)	The teacher demonstrates a belief in multiple models of achievement and works to differentiate across and within student groups.	The teacher demonstrates a belief in culturally relevant education and generally makes the classroom a site of academic success for all students.	The teacher demonstrates a belief in culturally relevant education and regularly co-constructs a learning environment where all students can be academically successful and culturally competent.

Attitudes toward Schools

While a teacher may be isolated to a single classroom, no classroom exists divorced from other classrooms or from its social context. A teacher must value the input and experience of others in the development of effective planning, instruction, and assessment.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
7(o) The teacher values planning as a collegial activity that takes into consideration the input of learners, colleagues, families, and the larger community.	The teacher expresses a value for only their own input in planning. (Note: rating of Unsatisfactory may trigger the creation of a deficiency report.)	The teacher expresses a value for multiple stakeholder voices in the act of planning and works to include those voices in instruction and assessment.	The teacher expresses a value for multiple stakeholder voices in the act of planning and generally includes those voices in instruction and assessment.	The teacher expresses a value for multiple stakeholder voices in the act of planning and regularly includes those voices in instruction and assessment.

Professional Dispositions

A disposition is an attributed characteristic that summarizes a trend of action in context. Dispositions congregate along four primary domains: Open-mindedness, Self-reflection, Curiosity, and Educational Equity.

Open-Mindedness Dispositions

The classroom is the first time many students are placed in close proximity to those from other backgrounds and cultures. In this space, a teacher must provide safety and challenge to all students, which requires a stance of open-mindedness to the ways and knowledges of diverse people.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
1(h) The teacher respects learners' differing strengths and needs and is committed to using this information to further each learner's development. 1(i) The teacher is committed to using learners' strengths as a basis for growth, and their misconceptions as opportunities for learning. 3(n) The teacher is committed to working with learners, colleagues, families, and communities to establish positive and supportive learning environments. 3(q) The teacher seeks to foster respectful	The teacher displays respect for only their own conceptions of a successful classroom environment and academic achievement. (Note: rating of Unsatisfactory may trigger the creation of a deficiency report.)	The teacher displays respect for multiple conceptions of a successful classroom environment and academic achievement and works to bring these to bear on the development of individual students.	The teacher displays respect for multiple conceptions of a successful classroom environment and academic achievement and generally brings these to bear on the development of individual students.	The teacher displays respect for multiple conceptions of a successful classroom environment and academic achievement and regularly brings these to bear on the development of individual students.

communication among all members of the learning community. 4(p) The teacher appreciates multiple perspectives within the discipline and facilitates learners' critical analysis of these perspectives. 7(n) The teacher respects learners' diverse strengths and needs and is committed to using this information to plan effective instruction. 8(p) The teacher is committed to deepening awareness and understanding the strengths and needs of diverse learners when planning and adjusting instruction.				
--	--	--	--	--

Self-Reflection Dispositions

The most important posture in a teacher's tool kit is self-reflection that drives changes and development in knowledge and action. Teaching is an art and a teacher must determine what a classroom needs, year after year, class after class. A teacher must use self-reflection as the foundation for effective planning, instruction, and assessment for all students.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
3(r) The teacher is a thoughtful and responsive listener and observer. 4(o) The teacher realizes that content knowledge is not a fixed body of facts but is complex, culturally situated, and ever evolving. S/he keeps abreast of new ideas and understandings in the field. 4(q) The teacher recognizes the potential of bias in his/her representation of the discipline and seeks to appropriately address problems of bias. 6(q) The teacher is committed to engaging learners actively in assessment processes and to developing each learner's capacity to review and communicate about their own progress and learning. 6(s) The teacher is committed to providing timely and effective descriptive feedback to learners on their progress. 6(t) The teacher is committed to using multiple types of assessment	The teacher displays a commitment only to their conception of the classroom and academic success. (Note: rating of Unsatisfactory may trigger the creation of a deficiency report.)	The teacher displays a commitment to receiving constructive feedback and works to make changes to practice based on that feedback.	The teacher displays a commitment to pursuing and receiving constructive feedback and generally makes changes to practice based on that feedback.	The teacher displays a commitment to pursuing and receiving constructive feedback and regularly makes changes based on that feedback toward a coherent, well-articulated pedagogy.

processes to support, verify, and document learning. 6(u) The teacher is committed to making accommodations in assessments and testing conditions, especially for learners with disabilities and language learning needs. 6(v) The teacher is committed to the ethical use of various assessments and assessment data to identify learner strengths and needs to promote learner growth. 9(m) The teacher is committed to deepening understanding of his/her own frames of reference (e.g., culture, gender, language, abilities, ways of knowing), the potential biases in these frames, and their impact on expectations for and relationships with learners and their families.				
--	--	--	--	--

Curiosity Dispositions

In common parlance, a good teacher is a lifelong learner. To this End, we look for a disposition toward curiosity, both toward understanding how things work but also toward imagining how things might be made better.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
3(p) The teacher is committed to supporting learners as they participate in decision-making, engage in exploration and invention, work collaboratively and independently, and engage in purposeful learning. 4(r) The teacher is committed to work toward each learner's mastery of disciplinary content and skills. 8(r) The teacher is committed to exploring how the use of new and emerging technologies can support and promote student learning. 9(n) The teacher sees him/herself as a learner, continuously seeking opportunities to draw upon current education policy and research as	The teacher displays a commitment only to their own current knowledge and does not feel the need to explore content or pedagogy further. (Note: rating of Unsatisfactory may trigger the creation of a deficiency report.)	The teacher demonstrates some exploration of content and pedagogy and displays a commitment to the posture of a lifelong learner. The teacher works to connect current and new learning in content and pedagogy to classroom practice.	The teacher demonstrates active exploration of content and pedagogy and displays a commitment to the posture of a lifelong learner. The teacher generally connects current and new learning in content and pedagogy to classroom practice.	The teacher demonstrates active exploration of content and pedagogy and displays a commitment to the posture of a lifelong learner. The teacher makes time to explore new content and pedagogical knowledge and regularly makes connections to their evolving classroom practice.

sources of analysis and reflection to improve practice.				
---	--	--	--	--

Educational Equity Dispositions

It is difficult to imagine that one teacher can change the world, but we look for teachers who are willing to try. While much can be accomplished within a single classroom, a strong educational equity and advocacy disposition also means forming alliances across classrooms and communities.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
2(m) The teacher respects learners as individuals with differing personal and family backgrounds and various skills, abilities, perspectives, talents, and interests. 2(n) The teacher makes learners feel valued and helps them learn to value each other. 5(q) The teacher is constantly exploring how to use disciplinary knowledge as a lens to address local and global issues.	The teacher displays the belief that the classroom and society are meritocratic and that academic success is only available to some students. The teacher expresses deficit language or attitudes. (Note: rating of Unsatisfactory may trigger the creation of a deficiency	The teacher displays respect for students as individuals and works to understand and include multiple cultures and funds of knowledge in classroom practice.	The teacher understands the classroom as a site for addressing local and global issues, displays respect for students as individuals, and generally includes multiple cultures and funds of knowledge in classroom practice.	The teacher uses the classroom as a site for addressing local and global issues, displays respect for students as individuals, and regularly includes all cultures and funds of knowledge in the classroom as part of daily practice.

Professional Competencies

Teaching competencies include proficiency in planning, teaching, and assessing, but there are other duties and responsibilities that transcend a learning segment. Professional competencies include self-presentation, self-representation, professional collegiality and demeanor, and also taking responsibility for those tasks entrusted to you.

Tag InTasc	Unsatisfactory	Developing	Proficient	Exemplary
1(j) The teacher takes responsibility for promoting learners' growth and development. 6(r) The teacher takes responsibility for aligning instruction and assessment with learning goals. 7(p) The teacher takes professional responsibility to use short- and long-term planning as a means of assuring student learning. 9(l) The teacher takes responsibility for student learning and uses ongoing analysis and reflection to improve planning and practice. 9(o) The teacher understands the expectations of the profession including codes of ethics, professional standards of practice, and relevant law and policy.	The teacher does not respect or appears ignorant of professional expectations. The teacher refuses to take responsibility for planning, instruction, or assessment. (Note: rating of Unsatisfactory will trigger the creation of a deficiency report.)	The teacher is knowledgeable of professional expectations and works to meet them regularly. The teacher works to take full responsibility for planning, instruction, and assessment.	The teacher is knowledgeable of professional expectations and meets them generally. The teacher generally takes full responsibility for planning, instruction, and assessment.	The teacher is knowledgeable of professional expectations and meets them regularly, even serving as a role model for other teachers. The teacher regularly takes full responsibility for planning, instruction, and assessment and seeks to maintain and elevate professional competencies.

(Brown & Laughter, Feb. 2020)

Vita

Lynnette Whitsitt Neu was born and raised in Duff, Indiana. After high school, she attended the University of Evansville in southern Indiana. Lynnette graduated summa cum laude from UE with a Bachelor of Science in Psychology and minors in Ethics and Sociology. In the fall of 2017, she entered the School Psychology program at the University of Tennessee, Knoxville. Lynnette graduated with an M.S. in Teacher Education with a concentration in Applied Behavior Analysis from the University of Tennessee in December of 2019. Following completion of a year-long internship with Schaumburg High School in Schaumburg, Illinois, part of the Illinois School Psychology Internship Consortium, she will receive her Ph.D. in School Psychology.