The Value of Teacher-Student Relationships: A Narrative Analysis of Teacher Evaluations in Washington State

Linnea Bostrom

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Education

University of Washington

2023

Committee:

Nancy Beadie

Shaneé Washington

Program Authorized to Offer Degree:

College of Education

University of Washington

Abstract

The Value of Teacher-Student Relationships: A Narrative Analysis of Teacher Evaluations in Washington State

Linnea Bostrom

Chair of the Supervisory Committee:

Nancy Beadie

College of Education

National and state initiatives to improve student academic outcomes in K-12 public schools have included efforts to reform teacher evaluation policies and practices. Prior research into teacher evaluation systems raises questions about how to accurately measure effective teaching practices. This study takes up these questions and uses Nel Noddings' (2005) concept of an "ethic of care" to identify the limitations of a variety of metrics used in research on teacher-student relationships. This study examines the three instructional frameworks used in Washington State through the lens of narrative analysis and finds that the language addressing teacher-student relationships neglects the unique context of individual classrooms and presents the teacher-student relationship as one of managing classroom procedures and student behavior. This study argues that without explicitly incorporating teacher-student relationships into these policies and frameworks, not only is this aspect of effective teaching left to individual interpretation, but teacher evaluation systems neglect the vital role that teachers play in supporting students' social, emotional, and academic growth.

TABLE OF CONTENTS

Introduction

Recent reforms to teacher evaluation systems have focused on what evidence is used to identify effective teachers, for example student performance on state testing or observations conducted by principals. Other attempts to improve the overall skill of the teaching force include proposals and suggestions to introduce monetary incentives to recognize effective teachers and motivate less-effective teachers. Two competing functions for teacher evaluation systems are present throughout the research on teacher evaluations; either teacher evaluations are a tool to exit ineffective teachers from classrooms or they are a tool to support professional growth and development. Most of the research does not question the foundation of teacher evaluations, however, and teacher evaluations are accepted as an important tool to improve academic outcomes for students. This project contributes to the current conversation surrounding teacher evaluations by raising questions about which aspects of teaching are included in, and therefor valued by, evidence-based instructional frameworks. Positive teacher-student interactions and relationships support student growth and academic outcomes, however teacher-student relationships are not directly named as an essential aspect of teaching in the instructional frameworks used to evaluate K-12 public school teachers in Washington State. Including teacher-student relationships in teacher evaluations recognizes the influence that the social and relational dynamics between teachers and their students have on student academic outcomes.

Research Question

The question driving this project is, how is the value of teacher-student relationships communicated through the language of teacher evaluation policies and instructional frameworks in Washington State? A supporting question is, what counts as evidence of teacher-student relationships in teacher evaluation documents?

Primary Sources

The primary sources in this project are Washington State legislative policies that outline the requirements for teacher evaluations (i.e., RCW 28A.405.100 and WAC 392-191A-060) as well as the three instructional frameworks used to evaluate teachers in Washington State, Danielson's Framework for Teaching, Marzano's Teacher Evaluation Model, and CEL 5D+ Teacher Evaluation Rubric 3.0. I use narrative analysis to identify evidence of what value is placed on relationship building in formal policy and teacher evaluation rubrics.

Narrative Analysis

Narrative analysis is an appropriate choice for this project because it frames the policies and instructional frameworks as elements of a story that communicate certain values and expectations from one person or group to another. In their chapter on a narrative approach to qualitative research Savin-Baden and Howell Major (2013) quote Hendry's (2010) perspective on narrative theory in which "narrative is not a method, but rather a process of meaning making that encompasses 3 major spheres of inquiry: the scientific (physical), the symbolic (human experience) and the sacred (metaphysical)" (Savin-Baden & Howell Major, p. 229). Analyzing teacher evaluations can be mapped onto this description of narrative analysis as follows: the process of identifying where teacher-student relationships show up in the policy and instructional frameworks is scientific inquiry; interpreting how teachers and principals make meaning of the instructional frameworks is symbolic inquiry; and identifying the values that are embedded within the policy language and instructional frameworks is sacred inquiry. Using a narrative approach to analyze these documents also acknowledges the temporal aspect of teaching that is often neglected by the teacher evaluation process. Teachers work with students over a school year yet rely on formal principal observations that only capture a minuscule part of the school year. Therefore,

teachers and principals must construct a narrative around a teacher's work with students that fills in the time surrounding the formal observations. The instructional frameworks support this narrative through a common language and set of terms and attributes that teachers and evaluators can use throughout the evaluation process.

The reliance on textual documents lends itself to a narrative analysis through examination of the structural form as well as the language used communicate the professional expectations for K-12 teachers in Washington State. According to Lisa M. Given in the entry about Narrative Analysis in the SAGE Encyclopedia of Qualitative Research Methods (2008), a structural form of narrative analysis "attends to how a story is composed to communicate particular communicative aims" (Given, p. 540). The primary source documents communicate the sequence of events and expectations for what instructional practices are indicative of high-quality teaching from policy makers to policy implementers (i.e., teachers and principals). A second component of the story told through the primary source documents is between evaluators and teachers as they interpret meaning from the documents to inform their work with students. Given (2008) states that in a narrative analysis "events are perceived as important, selected, organized, connected, and evaluated as meaningful for a particular listener" (p. 540). From this perspective of narrative analysis, the instructional practices that are presented in the policy documents communicate what elements of teaching are valued and how those elements should be measured during the evaluation process.

There are several factors to consider when analyzing the narrative of the instructional frameworks. First, the structure and organization of the instructional framework illuminates what elements of teaching are prioritized or highly valued. This includes looking at the domains (Danielson and Marzano) and dimensions (CEL 5D+) as well as the criteria within those domains.

Second, the language used to explain or describe the criteria and the performance levels can shed light on what the instructional framework is aiming to communicate to teachers and their evaluators. Third, the structure of the rubric that accompanies each criterion can also be used to interpret how teachers and evaluators might interact with or use the instructional framework as a tool for practice as well as evaluation. Each instructional framework has its own narrative and set of values that fits within the minimum criteria for K-12 public school teachers in Washington State and those values are communicated to teachers and principals both implicitly and explicitly when a district adopts one of the three instructional frameworks. Examining the selection process falls beyond the bounds of this project but could be an interesting course for future inquiry in qualitative, quantitative, or a mixed methods research study.

Positionality Statement

Throughout this project I kept noticing the tension I felt between how problems in education are presented in policy and education research and my own perspective on the nuances and challenges that teachers navigate every day based on six years of teaching in public school in Washington State. There were moments when I was reading the background research on teacher evaluations and teacher-student relationships that my own perspective and experiences clouded my understanding and interpretation of findings, my professional identity as a teacher is embedded throughout this project. My perspective is shaped by my experiences as a high school science teacher, and while it creates a bias in my analysis of the primary sources it is also valuable for my interpretation of these documents. The philosophy and values I hold in my identity as a teacher are not entirely separate from the philosophy and values I held during this project. Briefly, my philosophy is that it is necessary to recognize students as whole beings who are always learning from interactions and the environment around them, that a student's value is

not determined by their report card or their ability to follow rules, and that the content and skills I teach in science classes are important but are not the only things I am responsible for teaching. These values alone are not enough to identify whether I am an effective teacher, but neither are they unrelated to my effectiveness as a teacher.

Chapter One

Teacher Evaluations as Measures of Effective Teaching

In the following section I will review the existing literature that is focused on teacher evaluations and teacher evaluation systems that use observations as a measure of teacher effectiveness, a critical component of teacher evaluations in Washington State. The federal education policy, Race to the Top (2012), incentivized teacher accountability measures using student achievement data to identify effective, as well as ineffective, teachers. Three years later, the Every Student Success Act (ESSA) (2015) loosened federal guidelines, shifting power back to states and local districts to decide how to evaluate teacher performance. Prior to both pieces of federal education policy the Center for American Progress published a report in June of 2009 about teacher evaluation practices, specifically focused on how teacher evaluations inadequately reflect teacher quality. The report calls into question principals' ability to differentiate among teachers and provide feedback to improve instructional practices. The report raises questions about teacher evaluation systems that rely on just one type of assessment (i.e., observation, portfolio, value added measures (VAMs)) and recommends using multiple measures to capture more dimensions of effective teaching. This report emphasizes teacher evaluation systems as a path to "dismiss ineffective teachers" (Center for American Progress Report, p. 4) and does not identify how teacher evaluations could be a tool to support professional growth. The question of whether teacher evaluations are primarily to exit ineffective teachers from public schools or if they are primarily for professional growth and development continues to appear in studies throughout the 2010s. How did policy and the national conversation about the purpose of teacher

evaluations in 2009 shape the culture and practices within districts and schools over the next decade?

In partial answer to this question, Close et al. (2020) conducted an overview of teacher evaluation systems four years after ESSA (2015). This study identified four assessment strategies to evaluate teachers including, value added measures (VAMs), teacher-level observations, student surveys, and student learning objectives (SLOs) (p. 10). Close et al. (2020) reviewed the changes in states' teacher evaluation systems from 2012 to 2018 and, importantly, presented the perceived strengths and weaknesses of the post-ESSA teacher evaluation landscape, though with limited survey responses from states' departments of education. Local control and formative feedback were identified as strengths of the new teacher evaluation systems while the weaknesses included too much variety, limited capacity to support the variety of systems and communicate between the state and individual districts, and "the need for the language of official policies to reflect different attitudes towards teachers" (p. 17). This study shows how education policy from 2012 to 2016 created opportunities for states' teacher evaluation systems to shift from high-stakes, punitive accountability measures to more formative, feedback driven evaluation systems. Close et al. (2020) cite the Danielson and Marzano instructional frameworks as two such systems that emphasize formative feedback and break teaching down into essential subcomponents, both instructional frameworks are used in Washington State and will be discussed more in Chapter 2. Taking the minimal responses to the strengths and weaknesses of current teacher evaluation systems into account, there is still room to question whether the Danielson and Marzano instructional frameworks utilize language that (in)adequately reflect the philosophical changes to teacher evaluation post-ESSA, namely more cooperative and formative evaluation processes.

The following three studies look specifically at teacher evaluation ratings and how principals identify, or fail to identify, effective and ineffective teachers. Kraft and Gilmour (2017) revisit the 2009 study by the New Teacher Project that explored the discrepancy between perceived teacher effectiveness and formal teacher evaluation ratings, a similar problem addressed by the Center for American Progress report, also from 2009. This study, published in 2017, looked at yearly distributions of teacher evaluation ratings as national and state evaluation policies changed from the time of the initial report. The findings from the new study indicated that new evaluation systems, with more than two rating categories, had not led to increased differentiation among teachers. Kraft and Gilmour (2017) compared formal and informal evaluation ratings from a district case study to understand how principals were utilizing the new evaluation systems. As states and districts reformed their teacher evaluation policies there appeared to be a disconnect between the policy makers and the "will and capacity of local actors" which explained why the new evaluation systems were not more successful at differentiating among teachers than the previous systems (p. 235). The will and capacity of local actors, specifically principals, comes across in the qualitative data from the district case study, where time constraints were cited as one reason principals rated teachers' performance higher on formal evaluations. An additional factor for some principals was teachers' motivation and potential, noting how a low score might be counterproductive to teacher growth (p. 241). One commonality between time constraints and teachers' motivation and potential is the tension that principals experience when rating novice teachers and experienced teachers on the same framework. Novice teachers are still developing their professional skillset and therefore principals questioned whether it was fair to evaluate them in the same way as an experienced teacher. Kraft and Gilmour (2017) found that principals used low evaluation scores selectively to

manage their own workload as well as to avoid demotivating certain teachers due to low ratings. An additional challenge that principals shared was the risk of using low evaluation ratings to remove a teacher and run the risk of filling the position with a lower-quality teacher (p. 242). One implication of this study is to identify the purpose or goal of teacher evaluations. As Kraft and Gilmour (2017) recognized in their introduction some scholars view teacher evaluations as a tool to dismiss underperforming teachers while others view evaluations as a tool to support teachers' professional growth and development (p. 235). Meanwhile, the district case study sheds light on how local actors carry out state level policies and that, while the state and federal policies have changed, the evaluation and staffing decisions that principals make have not.

In a similar study Grissom and Loeb (2017) investigated how principals subjectively evaluate teachers in high- and low-stakes environments. This study defined the "high-stakes environment" as the ratings principals gave on formal evaluations, measured via the Instructional Performance Evaluation and Growth System (IPEGS), and "low-stakes environment" as the informal reports to researchers in structured interviews. Grissom and Loeb (2017) found that even when principals reported ineffective performance on essential teaching practices in a low-stakes setting they rarely gave low ratings on the formal, high-stakes, teaching evaluation (p. 388). Along with this finding Grissom and Loeb (2017) identified patterns in rating inflation and deflation, specifically instances of principals inflating ratings for beginning teachers, a finding that is supported by the qualitative evidence collected by Kraft and Gilmour (2017). Grissom and Loeb (2017) suggest "frequent and rigorous evidence gathering" as a way to collect "truer" ratings from principals that would result in more accurate feedback and support instructor growth (p. 389). This recommendation is challenged by Kraft and Gilmour's (2017) finding that time constraints and principals' workload are factors in formal evaluation ratings, however it does

align more with the philosophy that teacher evaluations are a tool for professional development rather than a tool to exit ineffective teachers from the system.

Looking more closely at observations as a measure of teacher effectiveness Steinberg and Garrett (2016) sought to understand what teacher observation scores measure, specifically looking at the relationship between student/class characteristics and the teacher's score on their evaluation. Steinberg and Garrett (2016) analyzed student achievement data and teacher evaluation data that were collected during the 2013 Measures of Effective Teaching Project (MET), a study conducted in six districts across the United States over a two-year period. Incoming student achievement was measured by states' ELA and math accountability tests and videos of teachers were evaluated using Danielson's Framework for Teaching. The overarching finding of this study is that teacher scores from classroom observations are significantly influenced by the context of the classroom (p. 295). The analysis found that teachers assigned higher achieving students received higher ratings in the observations, and that aspects of a teacher's instructional practice that were evaluated in an observation were influenced more by student achievement/characteristics than by the teacher's instructional ability (p. 307). This study is limited by the focus on incoming student achievement and does not discuss how other factors in classroom context (i.e., racial and ethnic identities, disability, and English language fluency) influenced the scores principals gave teachers based on observation. The implication here is for teacher evaluation systems that rely heavily on observations to include measures that are focused on teacher's ability in the context of their classroom rather than measures that misattribute class/student characteristics to teacher skill.

Teacher evaluation policy continues to shift as the United States, as well as individual states, focuses on improving student academic outcomes. Prior to Race to the Top (2012) and

ESSA (2015) the Center for American Progress published a report in 2009 that presented teacher evaluation systems as a tool to identify ineffective teachers and remove them from the public education system. Close et al. (2020) examined how teacher evaluation systems shifted in the post-ESSA policy landscape and recommended that teacher evaluations use multiple methods to assess teacher effectiveness. This study identified the lasting tension surrounding the purpose of teacher evaluations, as tools to identify ineffective teachers or tools to support professional growth. Kraft and Gilmour (2017) collected evidence from a case study that showed the disconnect between policy makers and local actors. Even with policy reforms principals continued to face time constraints that influenced how they approached teacher evaluations, citing their own workload, teacher motivation and potential, and staffing decisions as important considerations during the teacher evaluation process. Grissom and Loeb (2017) found that principals were more likely to use lower ratings for teachers in interviews that weren't related to a teacher's employment with the school district. This indicates that principals can identify ineffective teaching practices and supports Kraft and Gilmour's (2017) findings that other factors influence how principals use low ratings in formal evaluations. Finally, Steinberg and Garrett (2016) identify a link between teacher observation scores and student academic achievement, bringing classroom composition and context into conversation with tools that measure teacher effectiveness. The dynamics between teachers and students, and specifically how those dynamics influence student academic outcomes is examined in the next section.

Teacher-Student Relationships and Student Academic Outcomes

The following section provides an overview of the research surrounding the impact of teacher-student relationships and interactions on student academic outcomes. Researchers

approach, describe, and measure positive teacher-student relationships and interactions in different ways, therefore, I use Nel Noddings' (2005) ethic of care to provide structure and language that supports the conceptions of teacher-student relationships and teacher-student interactions used in the studies presented in this section. As Noddings (2005) defines and establishes what caring means in schools for teachers and students she states that "caring is a way of being in relation, not a set of specific behaviors" (p. 17). This relational approach to caring is counter to the evaluation tools discussed in the previous section, all of which attempt to specify and measure effective teaching practices (most commonly through observations and/or VAMs). The relational nature of teaching and learning requires an understanding that "who the teacher is, who the students are, what they are trying to accomplish separately and together all matter when designing instruction," and, consequently, in teacher evaluation policy and practice (Noddings, 2005, p. 8). Holding the identities of teacher and students, and their individual and shared goals, all together challenges the universalizability of teacher evaluations and the current tools used to identify effective teaching practices. Recall the finding from Steinberg and Garrett's (2016) study into teacher observation scores which showed that incoming student academic achievement had an impact on perceived teacher skill, specifically in aspects of classroom management and instruction (Steinberg & Garrett, 2016, p. 312). Accounting for the identities and goals of teachers and students in evaluation tools could result in a more reliable way to identify and recognize effective teachers and their practices.

The unique combination of teachers, students, and context (i.e., grade, subject, school, community, etc…) influence the classroom dynamic. Many of the studies presented in this section used student performance data from states' accountability tests, most commonly in ELA and Math, to identify the positive impact of teacher-student relationships on student academic

outcomes. This approach to teacher-student relationship research leverages the emphasis on student academic achievement data in education policy to justify why teacher-student relationships matter in K-12 schools. This communicates that teacher-student relationships are only important *if* there is a measurable increase in student academic achievement and accepts student academic achievement as an appropriate measure of teacher-student relationships. This presents an overarching tension in this body of literature that uses student participation and engagement as indicators of positive teacher-student relationships thus generalizing what positive teacher-student relationships look like in classroom and school settings. Attempts to define what a positive teacher-student relationship looks like will privilege certain values over others. Noddings' (2005) ethic of care offers responsiveness as an alternative to accountability, noting that "the most basic idea of relational caring is to respond to each individual in such a way that we establish and maintain caring relations" (p. xviii). This approach allows for flexibility while maintaining the expectation that teachers and schools fulfill their professional responsibilities to students, communities, and the broader society. Accountability in education is focused on finding or designing a single tool or policy that will successfully improve the school system while responsiveness is focused on supporting the unique needs of students in their specific context.

Cornelius-White's (2007) meta-analysis of learner-centered education looks at the impact of teacher-student relationships on student cognitive and behavioral outcomes. Drawing from Carl Roger's (1969) classical model as well as the current model of learner centered education Cornelius-White (2007) brings the relational nature of teaching and learning into focus. The classical model emphasizes "the perception of care by the student" as well as "an initial genuine trust in learners by the facilitator" (p. 114) both of which contribute to creating a flexible and

collaborative learning environment. Relational practices that fit a learner-centered model of education include "honoring student voices, adapting to individual and cultural differences, encouraging learning, thinking, and having learner-centered beliefs" (p. 115). In addition to the five practices just listed, genuineness, positive relationships, and non-directivity are included in the list of codes for independent (teacher) variables (p. 118). Cornelius-White (2007) identified correlation as the "conservative and most appropriate primary statistic to represent the findings" which include positive correlations for the impact of learner-centered practices on student cognitive and behavioral outcomes (p. 117). Citing Fraser, Wahlberg, Welch, and Hattie's (1987) assertion that a "correlation great than r=.20 is 'well worth pursuing'" a correlation of r=.45 for students 'critical and creative thinking should therefore be well worth pursuing (p. 131). Cornelius-White (2007) notes that critical and creative thinking is often measured through observations of occurrence; while this metric is not easily translated into other more common measures of student academic outcomes (i.e., state accountability tests) it is still a significant finding for the benefits of learner-centered teaching practices. In addition to the cognitive outcomes Cornelius-White (2007) also found several behavioral outcomes for learner-centered teaching practices, highlighting "large increases in participation/initiation (r=.55), satisfaction (r=.44), and motivation to learn (r=.32)" (p. 131). Identifying cognitive and behavioral outcomes acknowledges that what students learn in school is not limited to the curriculum and while state accountability tests are presented as cognitive measures, they also incorporate behavioral components that often go undiscussed, like a student's ability to sit quietly at a desk for an extended period of time.

In a longitudinal study on the influence of peer status and teacher-student relationships on behavioral engagement, Engles et al. (2016) expanded on existing literature to better understand

how classroom social dynamics impact student engagement. Student engagement serves as a proxy for academic achievement in this study because it is "an important factor in adolescents' educational outcomes" (p. 1192). Student participation data served as the measure of behavioral engagement and students were grouped based on peer perceptions of who had a positive or negative relationship with the teacher (p. 1192). This study characterized positive teacher-student relationships as "warm, sensitive and responsive interactions" while negative teacher-student relationships have "high levels of teacher conflict and lack of security" (p. 1193). The findings from this study add to the landscape of teacher-student relationship research by utilizing multiple observers to collect evidence of positive and negative teacher-student relationships. Engles et al. (2016) used the same questionnaire over three years to collect student self-reported data on their behavioral engagement. The questionnaire included items like "I pay attention in class" and "I work hard on my schoolwork" (p. 1196). The researchers used peer nomination to identify students with positive teacher-student relationships as well as students with negative teacher-student relationships to see how behavioral engagement differed for the two groups of nominees (p. 1197). Engles et al. (2016) found that negative (high teacher conflict) teacher-student relationships "consistently predicted relative decreases in behavioral engagement" while positive (warm, sensitive, and responsive) teacher-student relationships showed that "behavioral engagement was stable over time" (p. 1201). This finding shows that the social the dynamics present in a classroom can influence student academic outcomes through behavioral (dis)engagement. Positive teacher-student relationships support behavioral engagement, one aspect of academic engagement, and those relationships depend on classroom social dynamics which connects back to Noddings' assertion that who the teachers and students are matters.

Shifting from the influence of teacher-student relationships on student engagement to the impact of teacher-student relationships on students' academic success, Split et al. (2012) looked at the stability and change of teacher-student relationships in elementary school over six years. Split et al. (2012) used conflict and warmth as the descriptors for teacher-student relationships and tracked the changes in teacher-student relationship quality (TSRQ) as well as student academic growth. This study recognized the influence of context, "characteristics of the teacher as well as broader aspects … (e.g., number of students, social-emotional climate)" on TSRQ, citing prior research that examined gender differences, ethnic differences, and income group differences on social interactions in classrooms (p. 1181). A key finding from this study was the pattern of less academic growth for students who had increasing conflict over time, and Split et al. (2012) identified "African American children, economically disadvantaged children and children with low IQ or initial behavior problems," as subgroups of students who were generally more at risk for academic failure (p. 1189). This finding is significant for teacher evaluations that measure teachers' classroom management without clearly describing what classroom management entails. Similar to how state accountability tests are presented as cognitive measures but also covertly serve as behavioral measures, classroom management is presented as a quality of effective teaching but also covertly relies on students' classroom behavior matching the expectations of the observer. Teacher evaluation systems do not outrightly name and recognize the impact teacher-student relationships have on student academic outcomes and therefore privilege certain conceptions of classroom management over all others, resulting in the dominant value-set being seen as natural, or universal.

Building from research focused on the impact of teacher-student relationships on academic achievement Allen et al. (2013) conducted a study examining how teacher

characteristics and student characteristics influence teachers' observation ratings. Their study sought to understand when teacher-student interactions "predict future achievement" and when those interactions "reflect preexisting student characteristics" in secondary classroom settings (p. 80). Allen et al. (2013) organized observable classroom practices and processes into three domains: emotional support, classroom organization, and instructional support. Evidence of the type and quality of teacher-student interactions came from videos of lessons from early in the schoolyear and Virginia's state accountability tests provided student achievement data. Allen et al. (2013) found that all three of the domains of teacher-student interactions could be used to predict future student achievement, however, instructional support and classroom organization "are likely to reflect both teacher skill and student background" (p. 91). The significance of this finding regarding teacher evaluations, which are used to identify effective teaching practices, is that student characteristics can be misattributed to teacher skill. An additional finding was that teacher-student interactions that reflected emotional support predicted future student achievement with "relative independence" from student background characteristics (p. 91). The implication being that teacher-student interactions that fall into the Emotional Support domain reflect the qualities and skills of individual teachers without the same dependence on student-characteristics. In a descriptive analysis of the types of teacher-student interactions categorized in the Emotional Support domain Allen et al. (2013) list "teacher's ability to establish a positive emotional climate, their sensitivity to student needs" and "regard for adolescent perspectives," all as predictors of higher relative student achievement (p. 92). This study presents empirical data showing the positive impact of teacher-student interactions on student academic achievement while also revealing that some interactions are dependent on both teacher and student characteristics.

In an earlier study that focused on teacher-student interactions and student engagement, Allen et al. (2011) reported on an interaction-based approach to improving instruction and student achievement in secondary schools, the My Teaching Partner-Secondary program (MTP-S). The program was developed around the domains of the Classroom Assessment Scoring System-Secondary (CLASS-S) and was designed to "enhance the fit between teacher-student interactions and adolescents 'developmental, intellectual, and social needs" while aligning to elements of high-quality teaching (p. 1035). This study measured student achievement in math and ELA, as well as additional subjects, with data from Virginia's state accountability tests. The study found that student achievement improved regardless of the content area of instruction, suggesting that teaching "skillfully also involves successfully relating to and interacting with students" (p. 1036). The researchers did find gains in student achievement; however, the gains were observed in the post intervention year, indicating that "student gains in achievement would occur only after teachers had the benefit of a year's worth of their own growth" (p. 1036). This study demonstrates that changes or adjustments to the learning environment does not lead to rapid student achievement gains, which has implications for the pace at which instructional interventions that come out of teacher evaluations can have observable effects.

Another important aspect to consider in relation to the context of a classroom is the organization of the school itself, from the academic program to the day-to-day aspects like the bell schedule. Schools with a programmatic focus on STEM or the arts operate within the larger structure of public education while offering significantly different student experiences compared with comprehensive or neighborhood schools. One programmatic difference that poses challenges to teacher evaluation tools is Project Based Learning (PBL) which is being incorporated into more learning environments in a variety of ways. Morrison et al.'s (2019) study

at a project-based STEM high school highlights a disconnect between teacher evaluation tools like observations and innovations and changes in learning environments designed to improve student academic achievement. Their study explored how teachers at the project-based STEM school designed the learning environment to both support and challenge their students who only had one class for the entire semester and worked on small or large projects connected to the learning objectives for the specific course (p. 1107). Morrison et al. (2019) found that the high level of contact between teachers and students, due to the structure of the academic day and school year, facilitated close relationships that communicated care and support from teachers to students. Through interviews and surveys students shared that "teachers cared about them and their work and were passionate about what the students themselves were doing," which speaks to the supportive environment teachers created and maintained (p. 1116). Along with the unique daily schedule the PBL curriculum emphasized open-ended questions to spark student-driven inquiry which is different from a class that relies on a "traditional" curriculum where students complete the same, or very similar, assignments. Morrison et al. (2019) discuss how teachers navigated the challenges that arose in a project-based environment, one of which was "off-task" behavior. One researcher observed that students looked "off-task" yet were able to articulately communicate their project when asked. This demonstrates that there are values associated with on-task and off-task behavior in school settings, and students who do not appear to be actively engaged are labeled as off-task. One teacher commented to researchers that "[off-task behavior] will always be a problem in a place like this," and yet Morrison et al. (2019) state that "off-task behavior and down time seem to be a necessary part of the learning process and development of twenty-first century competencies" (p. 1112). The comment from the teacher illustrates the tension surrounding PBL environments and expected student behaviors. The student experience

at this STEM school was significantly different due to the structure of the school day and the academic program from what the current generation of teachers and researchers experienced as high school students. Here, the implication is for teacher evaluations that rely on observations and the implicit value that the observer places on "off-task" behavior. If off-task behavior and down time are necessary but counter to the implicit value-set, what does that mean for teacher evaluation-ratings?

Descriptions for teacher-student relationships and interactions varied across studies, Allen et al. (2013) used three categories of interactions, emotional support, classroom organization, and instructional support while Split et al. (2012) used warmth and conflict to describe positive and negative teacher-student relationships respectively. Engles et al. (2016) had students nominate peers for both positive and negative teacher-student relationships based on perception who which student did (or did not) get along well with the teacher (Engles, 1197). Morrison et al. (2019) cite Cornelius-White's (2007) meta-analysis of learner-centered education and identified four of the nine associated teacher variables, "empathy, warmth, nondirectivity, and encouragement" when discussing teacher-student relationships at the project-based STEM school (Morrison et al., 1106). Regarding academic achievement, two separate articles by Allen et al. (2011, 2013) used Virginia's state accountability tests, Engles et al. (2016) used self-reported data on students' behavioral engagement, and Split et al. (2012) used the Woodcock-Johnson III Tests of Achievement. While these studies show that teacher-student interactions and relationships have an impact on student academic outcomes they also illustrate how challenging it is to standardize or universalize what constitutes a positive teacher-student relationship. The context of the classroom, not only the demographic makeup of the teacher and students but also the culture and climate of the learning environment, matters for student achievement.

Challenges Associated with Measuring Teacher-Student Relationships

In the next chapter I present my analysis of the instructional frameworks that are used in Washington State to evaluate teachers. Considering that a goal of teacher evaluation systems is to identify (in)effective teachers, I am particularly interested in the language the three instructional frameworks use to define and describe effective teacher-student interactions and relationships. Noddings' (2005) discussion of the ethic of care and caring relations creates space to challenge how instructional framework present the role of teacher-student relationships in an effective teaching practice. Additionally, because the instructional frameworks in Washington state rely on observations, students and their behavior are also being observed and evaluated. Here, Nodding's (2005) assertion that "caring is a way of being in relation, not a set of specific behaviors" (p. 17) provides an opportunity to challenge the accepted (expected) set of student behaviors in a positive learning environment. Before introducing the analysis of the instructional frameworks, there are two propositions from Split et al.'s (2012) discussion of their study that provide an example of how a caring relation can challenge a perception of teacher-student relationships that is present in educational research. The two propositions are related, the first is that warmth is a dyadic characteristic of teacher-student relationships and second, is that conflict is a child-driven characteristic of teacher-student relationships. There are also two limitations to consider for these propositions which are that conflict and warmth were not defined by Split et al. (2012) and that only teacher reports were used to identify teacher-student relationship quality (TSRQ). Support for the second proposition, that conflict is child-driven, was found in the "moderately high year-to-year stability" for the level of conflict between teachers and students, which is contrasted with the low year-to-year stability found for warmth (Split et al., 2012, p. 1190). The proposition for conflict in teacher-student relationships, while statistically supported

in this study, goes against Noddings' (2005) ethic of care in which a caring relation requires that "both parties must contribute in characteristic ways" (p. 16). Arguing that conflict is a child-driven characteristic of TSRQ ignores the teacher's contribution to the teacher-student relationship and disregards the power dynamics between teachers and students, which could potentially absolve teachers of their responsibility to support the social and academic growth of the student.

The language of the proposition can be interpreted in such a way that the students are responsible for the conflict in the relationship with their teacher without needing to consider the broader context of the learning environment and the social dynamics that exist between teachers and students. This shift in attention can result in the student being blamed for the amount of conflict in the teacher-student relationship and consequently for their own diminished academic growth because conflict leads to being academically disadvantaged. This is a problematic interpretation considering not only the age of the students in this study (elementary school), but also the demographics of the associated subgroups of students identified in the study. The study found that "children with a low IQ, early behavior problems, African American membership, or disadvantaged SES" were on relationship trajectories that put them at increased risk for future academic failure (Split et al., 2012, p. 1190). Considering that any evaluation of a teacher is also an evaluation of student behavior, this interpretation has serious implications for measuring teacher-student relationships in formal evaluations.

It is unrealistic to expect classrooms to be free of social conflicts between students and between teachers and students and recognizing that teachers can cause or initiate conflict with students is an important aspect of evaluating teacher-student relationships. Teacher-student relationships are an essential, yet complex, aspect of teaching and learning. In comparison with

the previous interpretation, consider the earlier study that Split et al. (2012) reference in the proposition about conflict which says, "conflict in the kindergarten teacher-child relationship may reflect, in part, child-driven effects on teachers' interpretations of relationship" (Silver et al., 2005, p. 54). Silver et al. (2005) acknowledged the role of the teacher in their discussion in a way that Split et al. (2012) did not. Specifically, Silver et al. (2005) present conflict as the teachers' interpretation of the teacher-student relationship which supports Noddings' (2005) description of a caring relation, in which both care-er and cared-for contribute to the relationship. Silver et al. (2005) go on to note that reports of closeness may be more reflective of teacher skill and that "these supportive interactions may be especially important for children exhibiting initially high levels of externalizing behavior" (p. 54). This second interpretation provides an opening to question why certain groups of students might be experiencing higher levels of conflict with their teachers because it recognizes that social (power) dynamics exist between teachers and students. Additionally, Silver et al. (2005) articulated the significance of teacher-student relationships for future student academic success. Any observation of a teacher is also an observation of the students in the classroom, and if certain student behaviors are implicitly used to identify effective teachers, then students who experience conflict in their relationship with teachers will continue to be academically disadvantaged. While the goal to eliminate all conflict from teacher-student interactions is unrealistic, conflict can be an indication that, as Noddings (2005) puts it, the teacher is "unable to make the connections that would complete the caring relation with their students" (p. 2). The implication here is that students across K-12 schooling are at different developmental stages, and just as novice and veteran teachers are currently evaluated by the same tool, teachers working with early elementary and late high school students are also evaluated by the same tool. There is an existing tension between establishing a teacher

evaluation system that is usable in all K-12 public schools and providing teachers and evaluators with detailed performance descriptions for what effective teaching looks like in a wide variety of classroom contexts.

Chapter Two

Teacher Evaluation Policy in Washington State

The organization of the policies and procedures for teacher evaluations in Washington State includes elements of the Revised Code of Washington (RCW), the Washington Administrative Code (WAC), and documents provided on the website for the Office of the Superintendent of Public Instruction (OSPI). The Washington State Legislature provides the specific criteria for teacher evaluations in Title 28a, Common School Provisions, of the Revised Code of Washington (RCW). RCW 28a.405.100 establishes the minimum evaluation criteria and provides the guidelines for district implementation of the evaluation system for certificated employees (i.e. classroom teachers) in public schools. Subsection (2b) lists the eight minimum evaluation criteria required by the state, and subsection (2e) establishes that the Superintended of Public Instruction "shall identify up to three preferred instructional frameworks … [that are] research-based and establish definitions or rubrics for each of the four performance ratings for each evaluation criteria" (WA Legis. RCW 28a, 2019). Subsection (2e) also states that each school district must adopt one of the preferred instructional frameworks. There are two significant aspects of this organizational structure for teacher evaluation policy; first, that school districts in Washington State choose which of three instructional frameworks to adopt and second, that these three frameworks are responsible for providing definitions and rubrics that align with the eight state minimum evaluation criteria. This creates room to examine how each instructional framework places value on teacher student relationship building because the instructional frameworks are not designed specifically for Washington State's teacher evaluation criteria. Each instructional framework is responsible for defining performance levels for the

essential aspects of teaching; therefore the role and importance of teacher-student relationships and interactions will be communicated differently by each instructional framework.

The Washington Administrative Code (WAC) provides brief descriptors for the eight minimum evaluation criteria that build on RCW 28a.405.100. OSPI cites WAC 392-191A-060 on the public facing document that presents the eight minimum evaluation criteria with corresponding descriptors. Listed in numerical order the minimum criteria are, (1) centering instruction on high expectations for student achievement, (2) demonstrating effective teaching practices, (3) recognizing individual student learning needs and developing strategies to address those needs, (4) providing clear and intentional focus on subject matter content and curriculum, (5) fostering and managing a safe, positive learning environment, (6) using multiple student data elements to modify instruction and improve student learning, (7) communicating and collaborating with families and school community, and (8) exhibiting collaborative and collegial practices focused on improving instructional practice and student learning (WAC 392-191A-060). Each preferred instructional framework is aligned to these eight state criteria and therefore school districts in Washington State can select which tool to adopt.

State Evaluation Criteria and Teacher-Student Relationships

While student learning is the focus of all eight criteria there are three (criterion 3, 7, 5) that are most relevant to the role and function of teacher-student relationships and interactions in teacher evaluations. The language used in criteria 3, 7, and 5 either requires teachers to work with students on an individual level or to recognize students' identities beyond their academic achievement. The descriptor for Criterion 3 states, "the teacher acquires and uses specific knowledge about students' cultural, individual, intellectual, and social development and uses that

knowledge to adjust their practice by employing strategies that advance student learning" (OSPI, Teacher Evaluation Criteria and Descriptors). This language communicates a professional responsibility that a teacher must use knowledge of students' identities beyond the classroom to support student learning, and obtaining that knowledge requires some level of relationship building and interaction with students as individuals. Criterion 3 combines instructional skills, particularly lesson and assessment design, with the relational skills associated with the teacher-student interactions that occur during instruction.

Criterion 7 is focused on communication and collaboration with families and the school community. The descriptor states "the teacher communicates and collaborates with students, families, and all educational stakeholders in an ethical and professional manner to promote student learning" (OSPI, Teacher Evaluation Criteria and Descriptors). The language used in this criterion requires teachers to involve students' support systems in their learning and intellectual growth. The inclusion of "all educational stakeholders" acknowledges that families and communities play an integral role in students' learning, and successfully collaborating entails interacting and building relationships with students, their families, and the broader community.

The emphasis in Criterion 5 is on the learning environment, and the descriptor states that "the teacher fosters and manages a safe and inclusive learning environment that takes into account: Physical, emotional, and intellectual well-being of students" (OSPI, Teacher Evaluation Criteria and Descriptors). While it is possible to interpret that all students will have similar needs in terms of physical safety, differences in personality, life experiences, and social dynamics means each student will have varied needs when it comes to emotional and intellectual well-being. Regarding emotional and intellectual well-being, research surrounding teacher-student relationships and student academic outcomes shows that the social dynamics in a

classroom influence the learning environment and student academic growth (Allen et al., 2013; Split et al., 2012). Classroom management is a common term used to refer to the learning environment, and while Criterion 5 does not name teacher-student relationships and interactions in the descriptor, it is reasonable to interpret that teacher-student relationships are being evaluated under this criterion.

An important aspect of teacher-student relationships to keep in mind when analyzing the preferred instructional frameworks is, who is the subject of this evaluation? Findings from previous research of teacher observations show that students and teachers are both contributing to the learning environment and therefore are both being evaluated (Allen et al., 2013; Steinberg & Garrett, 2016). For this reason, analyzing the rubric language for the elements aligned with the State Criteria 3, 7, and 5 is necessary to identify who is the subject of the evaluation. Is the evidence of a teacher's ability to differentiate, collaborate with families and community, and foster a safe, positive learning environment based on student behavior? If so, to what extent is that reasonable and who is in the position to make the decision about what student behavior is evidence of the three professional expectations just listed?

Preferred Instructional Frameworks in Washington State

There are three preferred instructional frameworks in Washington State, CEL 5D+ Teacher Evaluation Rubric 3.0, Charlotte Danielson's Framework for Teaching (2011), and Marzano's Teacher Evaluation Model (OSPI Website). While all three instructional frameworks meet the minimum evaluation criteria and use a four-level rating (Unsatisfactory, Basic, Proficient, and Distinguished) the frameworks use slightly different domains and language to structure the comprehensive set of criteria encompassed by each individual framework.

CEL 5D+ Teacher Evaluation Rubric 3.0 was first published by the Center for Educational Leadership at the University of Washington in 2012 and updated in 2016. The CEL 5D+ instructional framework uses six dimensions to organize 30 elements, the six dimensions are: (1) Purpose, (2) Student Engagement, (3) Curriculum & Pedagogy, (4) Assessment for Student Learning, (5) Classroom Environment & Culture, and (6) Professional Collaboration & Communication. Each element within the six dimensions has its own rubric with performance descriptions. OSPI provides supporting documents for the CEL 5D+ instructional framework that include bulleted explanations and guiding questions for all dimensions except for Professional Collaboration & Communication.

Danielson's Framework for Teaching was first published by ASCD in 1996, followed by an updated edition in 2007. In 2011 an enhanced version of the Framework for Teaching was released. The enhancements were driven by the 2009 Measures of Effective Teaching (MET) study for which the Framework for Teaching was selected to evaluate teacher performance (Danielson, iv). Danielson's instructional framework uses four domains to organize 22 criteria, the four domains are: (1) Planning and Preparation, (2) The Classroom Environment, (3) Instruction, and (4) Professional Responsibilities. Each domain houses five to six criteria with each criterion having two to five associated attributes for a total of 76 attributes. Each criterion has a rubric with performance descriptions along the four-level rating system as well as "critical attributes" for each level that provides additional description or explanation for teaching practices at each performance level.

The third preferred instructional framework in Washington State is based on the New Art and Science of Teaching, written by Robert J Marzano which was first published by ASCD in 2007 with a second version published in 2017. OSPI refers to this instructional framework as "The

Marzano Teacher Evaluation Model," which uses four domains to organize 19 criteria; the four domains are (1) Classroom Strategies and Behaviors, (2) Planning and Preparing, (3) Reflecting on Teaching, and (4) Collegiality and Professionalism. Each domain contains anywhere from two to nine criteria and each criterion has between one and nine elements for a total of 60 elements.

The three instructional frameworks are designed to communicate the essential practices of effective teaching. In order to analyze the three frameworks, the "At A Glance" documents created by OSPI align the three unique taxonomies to the eight state criteria. Additionally, the "At a Glance" documents provide insight into how educational leaders in Washington State interpreted and determined how the instructional frameworks meet the Washington State teacher evaluation criteria. Each "At A Glance" document clearly communicate which elements, attributes, and/or criterion in the instructional framework are used to specifically evaluate State Criteria 3, 7, and 5.

State Criterion 3: Recognizing Individual Student Learning Needs

	CEL 5D+	Danielson	Marzano
State Criteria 3: **Recognizing individual student learning needs and developing strategies to address those needs**	SE2: Ownership of Learning SE 3: Capitalizing on students' strengths CP4: Differentiated instruction for students A4: Teacher use of formative assessments	1b: Demonstrating knowledge of students 3e: demonstrating flexibility and responsiveness	3.1: Effective scaffolding of information within a lesson 3.2: Planning and preparing for the needs of all students

CEL 5D+ Rubric for Instructional Growth and Teacher Evaluation

There are four CEL 5D+ criteria, from three domains of the framework (student engagement (SE), curriculum and pedagogy (CP), and assessment (A)) that are aligned with state

Criterion 3. The first CEL 5D+ criterion is Student Engagement 2 (SE2), ownership of learning, which the rubric describes as teacher-provided "opportunities and strategies for students to take ownership" (Center for Educational Leadership [CEL], 2016a, p. 3). The metric used to distinguish between performance level is whether "some" (Proficient level) or "most" (Distinguished level) of the "locus of control" is with the students (p. 3). The supplemental CEL 5D+ Teacher Evaluation Glossary defines ownership as "students having choice over the ways in which they learn and make meaning of complex concepts, actively engaging with content through these choices," but does not define locus of control (CEL, 2016b, p. 3). Criterion SE2, along with the definition of ownership are reminiscent of the learner-centered teaching practices discussed in Cornelius-White's (2016) meta-analysis, such as "explicit attempts to encourage higher order thinking and respect for divergent opinions" (p. 131) both of which had positive impacts on students' academic outcomes. However, the CEL 5D+ framework does not provide examples of what SE2 looks like in classroom settings so it is unclear if the framework views this criterion as a learner-centered teaching practice. Additionally, the wording of "teacher provides opportunities and strategies" raises accessibility questions: is it enough to provide opportunities or is a teacher also expected to support students in accessing those opportunities? Are the opportunities based on the teacher's knowledge of students' interests? The language in SE2 does not match the language of state Criterion 3, specifically there is no reference to "specific knowledge about students" present in SE2 when that is an emphasis in the descriptor for state Criterion 3. The opportunities for student ownership of learning look different depending on the grade level and/or subject area, and without additional explanation or description this "core element [of] quality instruction" is open to interpretation by teachers and principals (CEL, 2016a, p. 1).

The second CEL 5D+ criterion aligned with state Criterion 3 is Student Engagement 3 (SE3), capitalizing on students' strengths, with a performance-level determined by how much knowledge of students' interests a teacher has and whether use of that knowledge is connected to unit goals. Of note is the change in language that describes Basic-level performance as, "teacher has knowledge," and Proficient-level performance as, "teacher capitalizes on," (CEL, 2016a, p. 3) indicating that the knowledge of student strengths serves a purpose beyond informational value. The wording in SE3 does reflect the emphasis on students' cultural, intellectual, and individual development present in state Criterion 3, however the language of SE3 prioritizes academic, as opposed to social, learning because knowing and capitalizing on knowledge of student strengths is qualified by its use towards unit goals. Again, without any examples of what SE3 looks like in practice there is a question about how many students in the class a teacher should be able to do this for. Is the distinguished-level teacher "capitalizing on students' strengths" for every student? Is there any evaluative consideration made for which subgroups of students a teacher is (un)able to successfully capitalize on student strengths? Additionally, if SE3 is ultimately interested in connecting student strengths to unit goals, that neglects student growth and learning that is facilitated by teacher-student relationships but ultimately not directly connected to a state academic standard.

The third CEL 5D+ criterion aligned with state Criterion 3 is Curriculum and Pedagogy 4 (CP4), differentiated instruction for students. CP4 is evaluated based on the number of strategies (i.e., none, one, multiple) a teacher uses to support individual student learning. The language in the rubric indicates that the difference between a proficient- and distinguished-level teacher is the use of "targeted and flexible supports within the strategies" (p.4) though there is no additional explanation or definition for what targeted or flexible mean in this context of, or in relation to,

student learning. The final CEL 5D+ criterion aligned with state Criterion 3 is Assessment 4 (A4), teacher use of formative assessments. A4 evaluates how teachers utilize assessment data to inform instructional decisions (i.e., provide supports and modifications to students) and the wording raises the same question about which subgroups of students are (not) receiving supports and modifications. Similar to how CP4 only placed emphasis on individual students at the distinguished level, A4 also uses targeted feedback to individual students based on assessment data as an indicator of distinguished-level teaching. CP4 and A4 raise similar questions about how many of the students in a classroom a teacher is successfully able to differentiate instruction (CP4) and use formative assessment data (A4) to successfully recognize and develop strategies to support individual student learning needs. Additionally, the final question posed for SE3 is relevant here as well; what, if any, evaluative considerations are being made to determine which subgroups of students a teacher is (un)able to successfully differentiate for to support student learning?

Criterion 3 focuses on "recognizing individual student learning needs and developing strategies to address those needs" (OSPI, Teacher Evaluation Criteria and Descriptors). Combined, all four CEL 5D+ criterion aligned with state Criterion 3, address the emphasis on differentiation for student learning. However, the language across all four rubrics leaves room for interpretation regarding how many students a teacher is successfully differentiating for at the Proficient and Distinguished levels. SE3 is the only criterion that calls for teacher use of specific knowledge about students' identities, and even then, the language does not recognize the role that teacher-student interactions play in student academic achievement. The lack of examples or more detailed description for what these criteria look like in practice highlights how much room for interpretation exists in the evaluation of K-12 teachers working in districts that adopted the CEL 5D+ instructional framework.

Danielson's Framework for Teaching

There are two criteria in Danielson's Framework for Teaching aligned with State Criterion 3, the first of which is 1b, demonstrating knowledge of students. This criterion comes from the first domain of Danielson's Framework for Teaching, Planning and Preparation. Criterion 1b specifies five types of knowledge about students: (1) child and adolescent development, (2) the learning process, (3) students' skills, knowledge, and language proficiency, (4) students' interests and cultural heritage, and (5) students' special needs. These types of knowledge about students appear in the critical attributes as well as the rubric, where the distinction between levels focuses on how a teacher uses their knowledge of students to support student engagement and learning. For example, a critical attribute at the Basic-level is that "the teacher recognizes that children have different interests and cultural backgrounds but rarely draws on their contributions or differentiates materials to accommodate these differences," while the critical attribute at the Proficient level is "the teacher is well informed about students' cultural heritage and incorporates this knowledge in lesson planning" (The Danielson Group [Danielson], p. 8). Additionally, the critical attribute at the distinguished level is that "the teacher seeks out information about their cultural heritage from all students," which adds another action to the way the teacher engages and interacts with students. Seeking out information from all students about their identities outside of the classroom is indicative of some value being placed on teacher-student relationship building.

The second criterion from Danielson's Framework for Teaching that is aligned with state Criterion 3 is 3e, demonstrating responsiveness and flexibility. There are three components under 3e, (1) lesson adjustment, (2) response to students, and (3) persistence (p. 66). Each component has its own definition provided separately from the rubric, and lesson adjustment is defined as the ability to make "minor and (when needed) major adjustments to a lesson" with the objective of

supporting students who are having difficulty learning (p. 66). Response to students is defined as a teacher's ability to "capitalize" on an unexpected event and turn it into a teachable moment, while persistence describes a teacher who uses alternate approaches when students experience difficulty learning. The distinction made between levels in the rubric focus on the efficacy or success of a teacher to respond to student difficulties or unexpected events in order to support student learning. The Distinguished level specifies that a teacher makes adjustments to support individual students rather than groups of students and reaches out to other teachers or stakeholders to support student learning (p. 69). Just as the rubric and critical attributes for element 1b used specific, action-oriented language to describe each performance level the rubric and critical attributes for element 3b describe how a teacher is interacting and working with their students. For example, a critical attribute at the Basic-level is that the teacher "conveys a sense to students of their own responsibility for their learning but is uncertain about how to assist them," while a critical attribute at the Proficient level is that the teacher "conveys to students that [they have] other approaches to try when the students experience difficulty" (p. 68). The specific and direct language used in the rubric and critical attributes for criteria 1b and 3e not only breaks down the steps for building positive teacher-student relationships it also describes the differences between performance levels through specific teacher actions.

The Marzano Teacher Evaluation Model

The first component from the Marzano framework that is aligned to state Criterion 3 is component 3.1, effective scaffolding of information within a lesson. This component emphasizes instruction over teacher-student interactions and the one-sentence description does not specify that scaffolding is meant to support student learning at a subgroup or individual level. The list of

possible teacher evidence that is provided includes student choice and initiative but does not explicitly indicate that teachers should have knowledge of specific student learning needs and strengths (OSPI, 2022b, p. 16). Instead, the emphasis is on the organization and delivery of content within a lesson giving the impression that in the Marzano framework scaffolding is done before implementing the lesson and does not necessarily depend on teacher-student interactions during the lesson.

The second component that is aligned with state Criterion 3 is Component 3.2, "planning and preparing for the needs of all students" (p. 17). Individual student learning needs are explicitly referenced in the brief description, with further emphasis on ELL, special education, and "students who come from home environments that offer little support for schooling" (p. 17). The wording for the third subgroup of students places more value on the structure of formal schooling than on student learning but does acknowledge that aspects of students' lives outside of school can have an impact on their academic achievement. The list of possible teacher evidence includes multiple teacher actions that emphasize differentiation for English language learners (ELL students) and identifies an awareness of the "purpose for the intervention" which indicates that it is not just enough to use multiple intervention strategies for a large group of students but to understand how different interventions support different learning needs (p. 17). The language in the rubric for component 3.2 does not specify that differentiation needs to happen at the individual student level, and instead the emphasis is on supporting all students at a sub-population level. The direct language not only names multiple sub-populations whose learning needs are not met by general instructional methods, but also frames the successful use of interventions for all students in a sub-population as a distinction between Basic and Proficient scores.

State Criterion 7: Communicating and Collaborating with Families and Community

	CEL 5D+	Danielson	Marzano
State Criteria 7: **Communicating and collaborating with families and school community**	PCC2: Communication and collaboration with parents and guardians PCC3: Communication within the school community about student progress	4c: Communicating with families	7.1: Engaging in positive interactions with parents and the school community about courses, programs, and school events 7.2: Engaging in timely and professional interactions with parents and the school community

CEL 5D+ Rubric for Instructional Growth and Teacher Evaluation

The first CEL 5D+ criterion that is aligned with state Criterion 7 is Professional Collaboration and Communication 2 (PCC2) which is focused on evaluating the general communication about student progress and the goals of instruction with families and guardians. At the Basic, Proficient, and Distinguished level teachers are expected to communicate with all parents and guardians, with increased attention to aspects of communication such as language needs and two-way forms of communication (CEL 2016a, p. 7). At the distinguished level the rubric recognizes that families and guardians have a role in communicating with teachers, but prior to that there is no distinction made between successful and unsuccessful communication.

The second CEL 5D+ criterion that is aligned with state Criterion 7 is PCC3, communication within the school community about student progress. My interpretation of this criterion is that teachers are expected to access the appropriate school resources to support and celebrate student learning. Notably missing from the rubric for PCC3 is an explicit statement that teachers do this for all students, even though a similar statement is included in the rubric for PCC2. Teacher-student communication about successes and challenges is included in the description for distinguished level performance, which recognizes some value in positive teacher-student

relationships and interactions. However, the emphasis is on effective teaching practices and responsibilities that do not depend on teacher-student interactions or describe why this communication is an important part of effectively teaching all students.

Danielson Framework for Teaching

Danielson's Framework for Teaching has one criterion, 4c, aligned with state Criterion 7. There are three elements that make up 4c, (1) information about the instructional program, (2) information about individual students, and (3) engagement of families in the instructional program (Danielson, p. 80). One of the indicators for this component is that there is two-way communication, but a limitation of that indicator is that there are no additional criteria to distinguish between successful and unsuccessful implementation of communication. One challenge that teachers, and school systems, face on a regular basis is the disconnect between the culture and expectations of the school and the cultures and expectations of families when it comes to what, how, and when communication should happen. While the Danielson framework acknowledges challenges to family's involvement in their student's education the is no elaboration on the variety of ways that teachers could collaborate with families so that each one is able to access the opportunities to support their student despite constraints on their time and other resources.

Criterion 4c raises questions about how teacher communication practices are evaluated and what is evidence of proficient or distinguished level performance. Is providing one type of opportunity for families to be involved and included in the educational program enough to meet the critical attributes of criterion 4c? How have Learning Management Systems (LMS) changed the way families and students access their academic progress and learning? LMSs have

fundamentally changed how students and families can access information about academic progress, however in the rubric for 4c a critical attribute of the Basic level is "school or district-created materials about instructional program are sent home" and "teacher maintains school-required grade book but does little else to inform families about student progress" (p. 82). Are LMSs considered to only be a grade book despite their capacity to organize and communicate units- and courses-worth of information?

Additionally, what does meaningful collaboration and involvement look like for families, and what is reasonable to expect teachers to do to provide those opportunities? What is the expected level of effort, on the part of a teacher, to facilitate the two-way communication that is an indicator of component 4c? Is it fair to evaluate teachers on something like two-way communication when there is a variable that is out of their professional control? The critical attributes at the Distinguished level rely on students following through with communicating to their families in various ways which provides an example of teachers being evaluated on student behaviors. This is an area of the Danielson framework where there is not much supporting description or example evidence. Providing clear strategies or examples of culturally sensitive communication would disrupt the implicit values that privilege white middle-class conceptions of what, how, ad when school should communicate with (or to) families.

The Marzano Teacher Evaluation Model

The first component in the Marzano framework that is aligned with state Criterion 7 is Component 7.1 which outlines ways that a teacher engages in "positive interactions" with parents and the school community (OSPI 2022b, p. 29). The framework provides possible teacher evidence which includes communication through existing structures like PTSA and advisory groups as well

as using communication to involve families and the school community in other classroom and school events. All but one of the teacher actions included in the list could successfully foster relationships with families and community members who are already comfortable and feel included in their students' school experience. The final bullet point, "seeks to learn families' past experiences with school systems," is a strategy that could facilitate relationship building with families who have had negative experiences with school systems, but the type of involvement is still determined by the teacher and school (p. 30). The evidence, as currently written, privileges cultural values and expectations of the school without consideration for the ways that schools have historically excluded and marginalizes certain students, families, and communities. The framework emphasizes the timeliness and clarity of communication and collaboration between the teacher and families without indicating that "successful" communication is necessary to distinguish between performance levels (p. 30). Considering that teacher evaluations are conducted by principals, the teachers' communication is only being evaluated from the perspective of school employees, leaving the extent to which the communication is clear to families and larger school community unevaluated. While it is possible that teachers and evaluators incorporate family feedback regarding communication in the evaluation process the rubric does not set that as an expectation. Communication and collaboration are inherently relational, yet the language defining how teachers are evaluated on this criterion does not require the inclusion of family and community perspectives thereby privileging the perspective of teachers and principals and leaving any reflection about the efficacy of communication up to their interpretation of the criterion.

The second component from Marzano's framework that is aligned with state Criterion 7 is 7.2, "timely and professional interactions" with parents and the school community (p. 30). This component is also evaluated by people who have an insider perspective of the school system

without a set expectation of including feedback from families and the community in the evaluation process despite the focus being on communicating to and collaborating with families and the school community. The main difference between components 7.1 and 7.2 is what information is being communicated by the teacher; component 7.1 is focused on general happenings in classes and at the school while component 7.2 is focused on student progress. Teaching is a relationship-based profession and evaluating a teacher's ability to build, foster, and manage relationships with diverse groups of students, families, and school communities is challenging. The aspect of communicating and collaborating that isn't defined by the Marzano framework is what a teacher is expected to do when communication breaks down or isn't effective. This leads to a similar question that was generated by state Criterion 3, what process is in place to identify which subgroups of students and families a teacher is (un)successfully communicating and collaborating with?

State Criterion 5: Fostering and Managing a Safe, Positive Learning Environment

	CEL 5D+	Danielson	Marzano
State Criteria 5: **Fostering and managing a safe, positive learning environment**	CEC1: Classroom arrangement and resources CEC3: Use of learning time CEC4: Student Status CEC5: Norms for learning	2a: Creating an environment of respect and rapport 2c: Managing classroom procedures 2d: Managing student behavior 2e: Organizing physical space	5.1: Organizing the physical layout of the classroom 5.2: Reviewing expectations for rules and procedures 5.3: Demonstrating withitness 5.4: Applying consequences for lack of adherence to rules and procedures 5.5: Acknowledging adherence to rules and procedures 5.6: Displaying objectivity and control

CEL 5D+ Rubric for Instructional Growth and Teacher Evaluation

The first CEL 5D+ criterion aligned to state Criterion 5 is Classroom Environment and Culture 1 (CEC 1), which focuses on classroom arrangement and resources, specifically addressing the physical space in the classroom and students' physical safety (CEL 2016a, p. 6). The language in the rubric emphasizes academic learning without acknowledging that students' sense of belonging can support learning and academic achievement. The focus on academic scaffolding and connection to content overlooks the effort and work teachers do to create a welcoming and inclusive learning environment thus undervaluing that work in the evaluation process.

The second CEL 5D+ criterion aligned to state Criterion 5 is CEC 3, use of learning time, which is focused on instructional time as well student (mis)behavior. The rubric establishes that evidence for a distinguished-level teacher is that students "manage themselves, assist each other in managing behavior, or exhibit no misbehavior" (p. 6). Not only does this create an instance where student characteristics are being attributed to teacher performance, there is no further explanation about what constitutes student (mis)behavior, which privileges certain types of student behavior over others. The emphasis in CEC3 is on student behavior rather than on students' physical, emotional, or intellectual well-being and does not describe specific teacher actions that would foster and manage the student behavior described in the rubric.

The third CEL 5D+ criterion aligned with state Criterion 5 is CEC 4, student status. The language throughout the rubric for CEC4 focuses on the social dynamics that are present in a classroom, namely teacher-student and student-student relationships. The performance description for a proficient-level teacher states, "teacher and students demonstrate positive…relationships that foster students' well-being and develop their identity as learners" and that interactions indicate that "all are valued for their contributions" (p. 6). While the emphasis is on students' social and

intellectual well-being there is no additional description of what positive relationships look like or that students have identities other than 'learner' that are present in the learning environment and are also important for teachers to support. Positive teacher-student and student-student relationships will look different at the different levels of K-12 education, but the current wording leaves the work of defining positive relationships up to the teacher and evaluator. This element raises the question, how does the evaluator measure that all students are valued for their contributions? Part of this question can be answered by the supplemental glossary for the CEL 5D+ rubric, which states that when *all* is used it means "that a preponderance of evidence from the available data shows that the teacher includes all students" (CEL 2016b, p. 1). This leads to a new question about what data is available to an evaluator, and again, how this is measured by someone who isn't in the classroom on a consistent basis?

The fourth CEL 5D+ criterion aligned with state Criterion 5 is CEC5, norms for learning. The emphasis in CEC5 is on classroom norms that result in identifiable social interactions. The specific patterns of interaction that are named in CEC 5 are "risk-taking, collaboration, respect for divergent thinking and students' cultures" (CEL 2016a, p. 6). In order for a teacher to foster and manage a learning environment with those qualities the teacher needs to have knowledge of their students as individuals and not just as students in an academic learning environment. This criterion also raises the question, how this is measured or evaluated by someone who is not a consistent member of the classroom community?

A consequence of not having detailed descriptions for the CEL 5D+ criteria that align to state Criterion 5 is that not all teachers and evaluators will place the same value on the importance of positive teacher-student interactions and relationships. CEC1 identifies content related materials and resources as elements of the classroom that contribute to a safe and positive learning

environment without considering how the environment relates to students' sense of belonging which is also an important factor in student academic growth and achievement. Additionally, teachers who are less inclined to get to know their students can still meet the 'proficient' or 'distinguished' category despite not holding knowledge about their students that would support student academic and personal growth. Relationships take work and effort and not all teacher-student relationships will be identical and despite research showing that positive interactions with teachers have positive impacts on student academic achievement (Allen et al., 2013; Cornelius-White, 2016; Engles et al., 2016; Split et al., 2012) the CEL 5D+ framework does not communicate a professional expectation that teachers should build relationships with students.

Danielson's Framework for Teaching

Danielson's Framework for Teaching has four criteria aligned with state Criterion 5, all of which are housed in Domain 2: The Classroom Environment. The first criterion is 2a, creating an environment of respect and rapport, which is focused on both the words and actions used in teacher-student and student-student interactions. The language used in the rubric and critical attributes acknowledges the different developmental needs of students based on their grade and age. General care and respect will look different in an elementary school classroom and a high school classroom, but the framework establishes that teachers in all classroom settings are expected to know what is appropriate for the age group they are working with (Danielson, p. 28). The descriptions used in criterion 2a emphasize classroom interactions, for example, the description of a Distinguished level teacher includes "the net result of interactions is that of connections with students as individuals" (p. 31). This phrasing goes beyond a general explanation of "classroom management" and acknowledges the relational component of teacher-student

interactions. There is evidence in this framework that the positive, supportive relationships that teachers build with their students, as individuals and whole classes, are valued and recognized. The difference between Proficient and Distinguished is whether connections with students are superficial or demonstrate a genuine interest in students' lives beyond the classroom (p. 31). This raises a question about what type of evidence can be provided to show a more-than-superficial teacher-student relationship, but it is clear that at both levels positive teacher-student relationships are valued.

While the rubric and critical attributes value positive teacher-student relationships and outline the types of interactions that facilitate building those relationships, the possible examples promote certain values over others. For example, in the provided examples at the Distinguished level one item is "students hush classmates causing a distraction while the teacher or another student is speaking" (p. 29). This example could be at odds with the expectation that students "contribute to high levels of civil interactions between all members of the class" because hushing is not universally interpreted as a civil interaction by students and adults alike (p. 29). In this instance, the critique of the examples provided for this specific criterion is intended to point out an inconsistency and suggest that there could be other examples of students holding each other accountable for their behavior that focus on the culture and norms of the specific classroom rather than the cultural value of being quiet when other people are taking. How teachers and evaluators make sense of the indicators listed with element 2a, specifically "respectful talk and turn taking," is important to establish because the unspoken expectations at school might be in conflict with the expectations of a student's family and/or cultural background. Acknowledging the breadth of ways students can participate display respectful talk is one way to shift this criterion from evaluating teachers *and* students to only evaluating a teacher's professional skills. As currently written,

students' exhibiting a unexpected behavior could result in the student being punished for a lack of compliance if the teachers lacks of requisite (as established by state Criterion 3) knowledge of students' cultural identities and background. Reading through the rubrics, critical attributes, and indicators for each component of the framework illuminates the inconsistencies that implicitly promote certain value sets and expected school behaviors.

The second criterion from the Danielson framework that is aligned with state Criterion 5 is 2c, managing classroom procedures. Criterion 2c emphasizes four aspects of a positive learning environment that are highly procedural and minimally relational: (1) management of instructional groups, (2) management of transition, (3) management of materials and supplies, and (4) performance of non-instructional duties (Danielson, 36). Managing instructional groups requires teacher knowledge of students in order to create student groups that will be productive and maximize learning and instructional time while communicating that students (at all levels) should be able to transition between parts of a lesson quickly, devaluing the social interactions that can happen in those transitions and breaks from structured learnings. The language in this criterion neglects the benefits that social interactions have for student learning and how these interactions can occur during transition times. Additionally, the expectation that is established by this component does not acknowledge the scaffolding or active learning strategies that support students but might interfere with quick transitions, such as students needing to move around the classroom. The priority is instructional time, not student learning with the assumption that students are only learning during instructional time. Depending on how a teacher or principal interprets "instructional time" different classroom procedures will be evaluated using this framework. A rigid interpretation might result in the whole class moving between parts of a lesson at the same time despite some students needing additional time to understand or finish the previous activity, while

a flexible interpretation might result in a lesson that is structured so that students can move from one activity to another at their own pace.

The third criterion from the Danielson framework that is aligned with state Criterion 5 is 2d, managing student behavior. This criterion focuses on the strategies teachers use to prevent or address student misbehavior, and while there is some language that provides evidence of relationships in the rubrics and critical attributes the bulk of what is being evaluated is procedural. For example, in the description of a Proficient-level teacher the rubric states "the teacher monitors student behavior against established standards of conduct" (Danielson, p. 43). The focus here seems to be on the teacher's awareness of what is happening in the room rather than how their relationships with students facilitate a positive learning environment. The supporting paragraph for criterion 2d specifies that "the atmosphere must feel business like and productive, without being authoritarian," and goes on to establish what constitutes a "productive classroom" (p. 40). According to this framework, indicators of a productive classroom include clear standards of conduct, student knowledge and understanding of what is expected of them, and what students can expect while in the classroom, however terms like "business-like" elicit certain expectations. Similarly, the term productivity is being narrowly defined, or left to be interpreted by teachers and principals, which privileges the implicit values embedded in the school system. Morrison et al.'s (2019) study at the project-based STEM school provides an example of how adult conceptions of productivity show up in observations and fail to accurately capture student learning. Despite the role teacher-student relationships play when it comes to managing student (mis)behavior, that value is not communicated through the language in the rubric for criterion 2d. This component specifically states that the atmosphere of the classroom shouldn't be authoritarian but does not describe the difference between authoritarian and non-authoritarian teacher-student dynamics.

The fourth criterion from the Danielson framework that is aligned with state Criterion 5 is 2e, organizing the physical space. This component addresses the physical safety of students via the layout of the classroom, access to resources and materials, and a lack of environmental dangers (i.e., dangling wires or dangerous traffic patterns) (Danielson, p. 44). Overall, these four criteria from Danielson's Framework for Teaching emphasize teacher interactions with students that maintain dignity and respect but do not incorporate language that explicitly recognizes the role of teacher-student interactions regarding students' emotional and intellectual well-being. Instead, the emphasis is on the use of instructional time and managing student behavior. The rubrics for components 2a and 2c emphasize the role of teacher-student relationships in managing a positive learning environment, but 2e is the only component that addresses the safety of the learning environment, and only considers the physical safety of students due to the inanimate objects in the classroom. Critical attributes for Criterion 3e, presented earlier in the discussion of state Criterion 3, include teacher reflection on practice, but there are no similar critical attributes for any component aligned with state Criterion 5.

Students bring their identities, experiences, and interactions with them to school and in classrooms, and these can both support and hinder their ability to access information and resources. The State Criteria do not define emotional and intellectual well-being, but student safety isn't just physical safety, and emotional and intellectual harm can come in many forms regardless of the age of the students or the subject matter of the class. The criterion specifies certain aspects of a classroom environment that are essential but leaves out other essential aspects that address student wellbeing and factors inside and outside of school that make it more difficult for some students to contribute in positive ways to the learning environment on a regular basis. The Danielson framework places value on relationships but does not establish professional expectation that

teachers need to work with all students, especially students they find challenging to work with, and are responsible for supporting the academic, social, and personal growth of each student.

The Marzano Teacher Evaluation Model

The Marzano Teacher Evaluation Model has six components aligned with state Criterion 5 that encompass the physical space of the classroom, the rules and procedures in the classroom, and teacher characteristics or dispositions that support a safe and positive learning environment. Component 5.1, organizing the physical layout of the classroom, establishes that access to materials and wall decorations should "facilitate movement and focus on learning" (OSPI, 2022b, p. 21). Neither list of possible evidence includes language that specifies that knowledge of students or student identities need to be considered when evaluating this component. Acknowledgement of individual student needs appears in the performance description for the Distinguished level, which states that "the teacher adapts or creates new strategies to meet the specific needs of students for whom the typical application of strategies does not produce the desired effect" (p. 21). The focus of component 5.1 is on the physical layout so far as it ensures classroom safety and facilitates student learning, does not mention how the layout can support student interactions with each other and the teacher.

Three of the components from the Marzano framework that are aligned with state Criterion 5 are all related to and address the rules and procedures in the classroom. Component 5.2, reviewing expectations for rules and procedures, emphasizes the teacher's role in establishing and communicating classroom rules so that students understand them. While the emphasis is on establishing and maintaining the rules and procedures, one possible piece of teacher evidence is that the teacher "strives to create environments which values a variety of cultural approaches to

learning" (p. 22). This seems to establish an expectation that teachers should be able to recognize that classroom routines, procedures, and rules have associated value sets and that incorporating other cultural approaches to learning is an essential aspect of effective teaching. However, when looking at the rubric for this component, the language is vague and does not describe how a teacher could or should incorporate a variety of cultural approaches to learning. The Distinguished-level performance descriptions for components 5.2 and 5.1 are identical, raising the question, what are the benefits and drawbacks to having a performance description that can be used for a variety of components? How does this influence how teachers and evaluators understand the distinction between a proficient-level and distinguished-level performance on the components that are attempting to measure how a teacher fosters and maintains the learning environment?

Component 5.4 from the Marzano framework, applying consequences for lack of adherence to rules and procedures, is the second component relating to classroom rules. The list of possible teacher evidence ranges from "tells students to stop" to "deliberately employs asset-focused language when describing student actions and behavior" (OSPI, 2022b, p. 23). The responses illustrate two different teacher-student interactions, one that is more authoritarian while the second utilizes a more relational approach to address student behavior. There is a stark contrast between examples, ones emphasize consequences and the other emphasizes approaching students with curiosity and using "asset-focused language" (p. 23). Split et al. (2012) found that elementary aged students who experienced increasing or consistent levels of conflict in their interactions with teachers were more likely to have diminished academic growth which is a reason to try and minimize teacher-student interactions that a teacher interprets as conflict with the student (p. 1189). Setting the professional expectation that teachers approach students with curiosity disrupts the power dynamic in classrooms where teachers occupy a role of authority over acceptable

behavior and encourages teachers to adopt a relational stance when responding to student (mis)behavior. Articulating this expectation in the instructional frameworks would not only communicate the value placed on teacher-student relationships it also provides an opportunity for teacher-centered feedback on their work with students rather than a reflection on student compliance with rules and procedures.

Component 5.5 from the Marzano framework, acknowledging adherence to rules and procedures, is a strategy that fits in with the Positive Behavior Intervention System (PBIS) which encourages teachers to recognize positive student behavior more frequently than correcting or acknowledging when student behavior does not meet the rules or procedures. PBIS encourages teachers to get to know their students and use asset-focused language throughout teacher-student interactions, however the possible teacher evidence provided with component 5.5 is not entirely aligned with PBIS strategies. One of the suggestions for possible teacher evidence is the use of a token economy or certificate of merit, which can help students recognize when their behavior is supporting a safe and inclusive learning environment, while on the other hand, such systems can also demotivate students who aren't recognized by the teacher or don't feel a sense of belonging in the classroom (OSPI Marzano, p. 23).

The two other components from the Marzano framework that are aligned with state Criterion 5 both focus on teacher dispositions or characteristics that support the management of a safe and positive learning environment. Criterion 5.3, demonstrating with-it-ness, requires that a teacher is aware of what is happening in the classroom at all times (OSPI, 2022b, p. 22). This includes having a physical presence that moves throughout the classroom as well as being aware of the interactions that are occurring between students throughout the class period. This component is more focused on the "social and emotional tone…as an indicator of cultural responsiveness, as

opposed to the presence of artifacts or decorations" which acknowledges that interactions are an important aspect of culturally responsive teaching. However, the possible student evidence speaks less to a social dynamic that supports healthy relationships between students and their teacher and more to a dynamic of compliance and adherence to rules. Student recognition that the teacher is aware of their behavior" does not mean that the student knows their teacher cares about their well-being and academic achievement. (p. 22). This interpretation is supported by the language of "adherence to rules and procedures" used prominently in components 5.4 and 5.5. The emphasis is on student compliance to rules rather than the understanding that a positive relationship between teacher and students can result in fewer disruptions to learning for all students and therefore a positive and safe learning environment overall.

Component 5.6 from the Marzano framework, displaying objectivity and control, is another characteristic or disposition that a teacher should display to show that they are fostering and managing a safe, positive learning environment. The brief description for this component explicitly states how teachers should build positive relationships with students by "displaying objectivity and control" (OSPI, 2022b, p. 26). The language in this component places value on a teacher's ability to reflect on and monitor the social dynamics that are present in the classroom, for example, the teacher "regularly engages in re-framing their interpretations of student behavior/misbehavior through an understanding of diverse cultural norms and lived experiences" (p. 24). The list of possible teacher evidence provides specific examples of how a teacher can be aware of the social and emotional tone of the classroom and also establishes an expectation that teachers are self-aware and self-reflective.

One strong piece of evidence that this component places a value on teacher-student relationships is found in the list of possible teacher evidence for component 5.6, communicating

value through language that sets the expectation that the teacher knows student identities and uses relational strategies to interact and engage with students. The specific piece of evidence states that the teacher "demonstrates sensitivity and awareness of moments when topics, content, or experiences may cause students to experience negative feelings or responses (fight or flight) due to lived experiences related to racial or gender identity and works to mitigate these or leverage these intentionally" (OSPI, 2022b, p. 24). This establishes that teachers would need to know their students' identities beyond the context of the classroom to be aware of moments when students might experience negative feelings. Similarly, the use of 'sensitivity' indicates that it's not enough to intellectually know students' identities but to also to have an emotional understanding of their students' identities and how those identities show up in the context of a learning environment.

Looking at the two ends of the interpretive spectrum for Criterion 5, teachers who value and foster positive relationships with their students will be able to demonstrate their proficiency and skill with all six components. However, teachers who hold a more authoritarian view of their interactions with students can still demonstrate that they are fostering and managing a safe, positive learning environment well enough to not be scored as basic or unsatisfactory. Considering the learning needs, including social, emotional, physical, and intellectual, of all students is a tall task in classes with 30 students. However, the language of these components could be more clear and specific about how teachers can foster and manage a safe, positive learning environment for *all* students rather than just for the students who's cultural value set matches that of the teacher and/or the school culture. Teaching and learning depend on the social dynamics (i.e., interactions and relationships) between individuals in a classroom regardless of the age of the students. A result of not using professional expectations (i.e., instructional frameworks) that undervalue the role of positive teacher-student relationships is that the labor of teachers who develop positive teacher-

student relationships, as well as the positive impact of those relationships on student academic outcomes, goes unrecognized.

Summary

The descriptors for state Criteria 3, 7, and 5 indirectly acknowledge that teaching is relational and, at the very least, teachers will interact with their students as they support student academic learning. Criterion 3 combines elements of teaching that are focused on a teacher's ability to build relationships with students (i.e., acquires and uses specific knowledge about students' cultural, intellectual, and social development) and the teacher's instructional skills to (i.e., adjust their practice by employing strategies that advance student learning) (OSPI, Teacher Evaluation Criteria and Descriptors). Each framework emphasizes different aspects of a professional teaching practice in the alignment with state Criterion 3. The language in the four criteria from CEL 5D+ emphasize the instructional, rather than relational, elements of differentiation, leaving room for the teacher and evaluator to interpret what, how, and when the specific knowledge about students is relevant to the teachers' evaluation. Somewhat similarly, the Marzano Teacher Evaluation Model also emphasizes the instructional aspects of differentiation, such as the logical organization of material and providing instructional interventions for subgroups of students. The emphasis on instruction implies that specific knowledge about students is mostly useful when it is related to differentiation and scaffolding. Danielson's Framework for Teaching approaches state criterion 3 differently and specifies that teachers should have knowledge of students' developmental stage, students' interests and backgrounds as well as demonstrate flexibility and responsiveness (Danielson, p. 6). The amount of additional explanation or

supportive description varies from framework to framework, leaving room for interpretation by teachers and evaluators.

All three frameworks approached state Criterion 7 similarly, emphasizing communication from the school to families about student progress. The Marzano Teacher Evaluation Model includes "demonstrates awareness and sensitivity to social, cultural, and language backgrounds of families" as possible teacher evidence, though the rubric defines Proficient-level communication as "timely and professional" (OSPI, 2022b, p. 30). The CEL 5D+ framework also includes the language-needs of parents and guardians as a consideration teachers need to make in their communication with families. Two-way communication is mentioned in the CEL 5D+ framework at the Distinguished-level and in the description of the criterion from Danielson's Framework for Teaching, but the State expectation is that communication is focused on the academic program rather than other aspects of student growth and learning, such as social.

The Marzano Teacher Evaluation Model and Danielson's Framework for Teaching both emphasize student (mis)behavior and use of instructional time in the criteria aligned with state Criterion 5. Three of the six components in the Marzano Teacher Evaluation Model are focused on classroom rules and procedures, with one component specifically concerned with applying consequences for "lack of adherence" (OSPI, 2022a). Two of the criteria from Danielson's Framework for Teaching start with the verb 'managing,' which does not convey a relational approach to fostering a safe, positive learning environment (however, state Criterion 5 uses the term 'managing' as well). CEL 5D+ uses more neutral terms like 'student status' and 'learning routines,' without providing additional description to frame the expectations. Without clarification the expectations can be set by the adult without considering that student well-being, student identities as learners, and student (mis)behavior are not static or uniform entities in a classroom.

Taken together, the three frameworks illustrate the challenges present with creating and using a tool to identify effective teaching practices and evaluate teachers. The CEL 5D+ framework provides the least amount of supporting description for each of the criteria, leaving room for teachers, principals, and school districts to interpret and define what evidence will satisfy a specific criterion. On one hand this is useful considering the framework is used across all grades and subjects, however there are implicit understandings and meanings connected to some of the expectations that would benefit from further definition. For example, naming a variety of student strengths in Student Engagement 3 would set the expectation that teachers can, and need to, identify more than just academic strengths to support student learning. This could be a strategy to reduce the number of conflict-driven interactions and communicate to students that a teacher cares about their well-being beyond their academic achievement. The Marzano Teacher Evaluation Model provides example evidence along with the rubrics for each component. The example evidence removes some of the room for interpretation on the part of teachers and evaluators and also consistently uses language that focuses on teacher actions, rather than students, in the rubrics (i.e., the teacher demonstrates, the teacher behaves, the teacher communicates). This language means that it is less likely the evaluator will misattribute student characteristics for teacher skill, but the framework still privileges certain student behaviors and classroom procedures over others, for example the emphasis on classroom rules and procedures in the components aligned with state Criterion 5. Danielson's Framework for Teaching goes a step further by adding descriptive paragraphs for each criterion explaining why it is an essential part of effective teaching and what it could look like in practice. However, these descriptive paragraphs do not completely avoid privileging certain value-sets over others, but it is easier to identify the values communicated through the framework, like the example provided earlier of

“students hush each other” which values turn taking (i.e., not talking over someone else or interrupting) and is not a universal cultural value.

Conclusion

In 2012 the Washington State Legislature updated the state's teacher evaluation policy following the national trend that targeted teacher evaluation practices in an effort to improve student academic outcomes. Overall, the updated state evaluation criteria and instructional frameworks provide more in-depth descriptions of teaching practices that could support the professional growth and development of teachers. However, the extent to which professional growth is an outcome of teacher evaluations depends on how principals and teachers interpret the instructional frameworks and understand the essential aspects of teaching. The three instructional frameworks provide different levels of supporting description and examples for each component; Danielson's instructional framework provides the most supporting information, the CEL 5D+ instructional framework provides the least, and the Marzano instructional framework is somewhere in-between.

Most of the evidence for teacher-student relationships from each framework is found in the components that are aligned with state criterion 5, fostering and managing a safe, positive learning environment. The environment and culture in a classroom are highly dependent on the social dynamics that are developed over the course of the school year and each framework emphasizes different aspects of the learning environment. The CEL 5D+ rubric is the only framework that refers to "teacher-student relationships" in the components aligned with state criterion 5, however the framework fails to provide supporting definitions or clarification leaving teachers and principals to interpret and define what this looks like for themselves. There isn't one way to foster a safe and positive learning environment and the framework does not engage teachers and principals in a critical reflection about how students experience the learning environment. The Marzano framework uses language that encourages teachers to be aware of the

"social and emotional tone" and exhibit objectivity and control when interacting with students while simultaneous emphasizing the importance of rules and procedures in establishing and maintaining a safe and positive learning environment (OSPI, 2022b, p. 22). This framework recognizes the influence a teacher has on the classroom dynamic and at the same time communicates that student compliance with rules is an effective teaching practice. The Danielson framework most clearly articulates the significance of the classroom environment and why it is an essential aspect of teaching, however the framework also presents teachers as managers of classroom procedures and student behavior. The critical attributes that accompany the rubric for component 2a, creating an environment of respect and rapport, emphasize the degree to which interactions among students and with the teacher are respectful. This component has the potential to disrupt practices and policies in schools that implicitly define what student behavior is (un)acceptable if it incorporated culturally responsive language that broadens the understanding of respectful behavior to include more than white middle class expectations.

Despite the different emphases across the frameworks all three leave room for interpretation on the part of principals and teachers. Considering that these frameworks are designed to be used across K-12 schools in a variety of contexts (i.e., primary vs secondary, rural vs urban, Title I vs highly capable) a certain amount of flexibility is necessary because effective teaching practices will vary based on context. At the same time, the language does not place value on, or convey the significance of, teacher-student relationships as an aspect of effective teaching. Teachers and principals are guided by the instructional frameworks as they co-construct a narrative that defines and identifies effective teaching practices over the course of the school year. There are components of the frameworks that suggest that acknowledging students' identities is an essential aspect of effective teaching, however, the frameworks continue to

privilege practices and values focused on academic achievement instead of student growth. Situating teachers as classroom managers and arbiters of appropriate behavior undermines the places in the frameworks where the relational aspects of teaching are already included. The frameworks favor teacher practices that result in student compliance rather than recognizing how teachers support their students social, emotional, and academic growth.

Principals in Washington State collect a significant amount of evidence through observations though state policy only requires that principals conduct two formal observations throughout the school year for a total of 60 minutes (Minimum evaluation criteria for the evaluation of certificated employees, 3(a)). This poses a challenge regarding how to evaluate aspects of teaching that develop over a longer period of time or aspects that might not be observable during short and infrequent observations; like teacher-student relationships. The current structure of teacher evaluation policy and the language in the instructional frameworks prioritize student academic growth and neglect to measure, and only minimally recognize, the aspects of teaching that support student growth and development beyond academic achievement. This imbalance, in evaluation policy and practice, means that teachers who place value on building positive relationships with their students do so without recognition of the impact of those relationships on student outcomes and the labor that goes into those relationships.

The emphasis on the procedural, rather than relational, aspects of effective teaching raises the issue of whether teacher evaluations are for accountability or for professional growth and development. Looking at the overall structure of the instructional frameworks is one way to assess whether professional growth or professional accountability is the overall purpose of teacher evaluations, does the framework facilitate conversations about professional practice or is it (un)intentionally presenting items to be checked off a list? If teacher evaluations are about

professional growth and effective teaching then evaluation tools should measure procedural and relational aspects of a teacher's work with students. The instructional frameworks provide one tool to evaluate teachers and, in an attempt, to be flexible enough to use in a variety of school contexts the relational aspects of teaching are overshadowed by aspects that are simpler to objectively measure.

In the final chapter of her book, Nel Noddings addresses the challenge of measuring what has been accomplished in a program that emphasizes care in schools, "perhaps we should bracket the word *measure* and just ask how we might evaluate such a program" (p. 179). In the decade since Washington State updated its teacher evaluation policy districts have adopted instructional frameworks and are now incorporating student growth data into the evaluation process as a measure of effective teaching. In addition to a decade's worth of technological advancement in schools the COVID-19 pandemic disrupted the normal operating procedures and highlighted many of the practices in K-12 schooling that are outdated and fail to support student growth. The efforts to improve students' social-emotional learning and include more culturally relevant teaching practices shows an increased societal value in cultivating inclusive learning environments. At the same time, this societal value is not reflected in the current teacher evaluation policies and procedures. Without codifying teacher-student relationships in teacher evaluation policies and instructional frameworks the system is not supporting the professional growth and development of teachers who have a duty to support students socially, emotionally, and academically.

References

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the Classroom Assessment Scoring System-Secondary. School Psychology Review, 42(1), 76–98. https://doi.org/10.1080/02796015.2013.12087492

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. Science (American Association for the Advancement of Science), 333(6045), 1034–1037. https://doi.org/10.1126/science.1207998

Center for Educational Leadership. (2016a). 5D+ Rubric for instructional growth and teacher evaluation.

Center for Educational Leadership. (2016b). 5D+ Teacher evaluation glossary. https://www.k12.wa.us/sites/default/files/public/tpep/frameworks/cel/5D%2B%20Teacher%20Evaluation%20Glossary%204-3-171.pdf

Close, K., Amrein-Beardsley, A., & Collins, C. (2020). Putting teacher evaluation systems on the map: An overview of states' teacher evaluation systems post–Every Student Succeeds Act. Education Policy Analysis Archives, 28, 58–. https://doi.org/10.14507/epaa.28.5252

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. Review of Educational Research, 77(1), 113–143. https://doi.org/10.3102/003465430298563

Donaldson, M. L. (2009). (rep.). *So Long, Lake Wobegon? Using teacher evaluation to raise teacher quality* (pp. 1–21). Washington , DC: Center for American Progress.

Engels, M. C., Colpin, H., Van Leeuwen, K., Bijttebier, P., Van Den Noortgate, W., Claes, S., Goossens, L., & Verschueren, K. (2016). Behavioral engagement, peer status, and teacher–student relationships in adolescence: A longitudinal study on reciprocal influences. Journal of Youth and Adolescence, 45(6), 1192–1207. https://doi.org/10.1007/s10964-016-0414-5

Grissom, J. A., & Loeb, S. (2017). Assessing principals' assessments: Subjective evaluations of teacher effectiveness in low- and high-stakes environments. Education Finance and Policy, 12(3), 369–395. https://doi.org/10.1162/EDFP_a_00210

Kraft, M. A., & Gilmour, A. F. (2017). Revisiting the widget effect teacher: Evaluation reforms and the distribution of teacher effectiveness. Educational Researcher, 46(5), 234–249. https://doi.org/10.3102/0013189X17718797

Minimum criteria for the evaluation of certificated employees, WA Legis., RCW 28A.405.100 (2012 & rev. 2019). https://app.leg.wa.gov/RCW/default.aspx?cite=28A.405.100

Minimum evaluation criteria, WA Legis., WAC 392.191A-060 (2013). https://app.leg.wa.gov/WAC/default.aspx?cite=392-191A-060

Morrison, J., Frost, J., Gotch, C., McDuffie, A. R., Austin, B., & French, B. (2021). Teachers' role in students' learning at a project-based stem high school: Implications for teacher education. International Journal of Science and Mathematics Education, 19(6), 1103–1123. https://doi.org/10.1007/s10763-020-10108-3

Noddings, N. (2005). The challenge to care in schools: an alternative approach to education (2nd ed.). Teachers College Press.

OSPI (n.d.). Teacher evaluation criteria and Descriptors. https://www.k12.wa.us/sites/default/files/public/tpep/frameworks/Teacher%20Criteria%20and%20Descriptors_ADA.pdf

OSPI (2018). CEL 5D+ Teacher evaluation rubric 3.0 at a glance. https://www.k12.wa.us/sites/default/files/public/tpep/frameworks/cel/CEL%205D%2B%20At%20a%20Glance.pdf

OSPI (2018). Danielson's framework for teaching at a glance. https://www.k12.wa.us/sites/default/files/public/tpep/frameworks/danielson/Danielson%20%282011%29%20At%20a%20Glance.pdf

OSPI (2022). The Marzano teacher evaluation model at a glance. https://www.k12.wa.us/sites/default/files/public/tpep/frameworks/marzano/Marzano%20at%20a%20Glance%202022.pdf

OSPI (2022). The Marzano teacher evaluation model by Washington state criteria. https://www.k12.wa.us/sites/default/files/public/tpep/frameworks/marzano/Marzano%20Rubrics%20with%20scales%2C%20potential%20evidence%20and%20elements%202022.pdf

Silver, R. B., Measelle, J. R., Armstrong, J. M., & Essex, M. J. (2005). Trajectories of classroom externalizing behavior: Contributions of child characteristics, family characteristics, and the teacher–child relationship during the school transition. Journal of School Psychology, 43(1), 39–60. https://doi.org/10.1016/j.jsp.2004.11.003

Steinberg, M. P., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? Educational Evaluation and Policy Analysis, 38(2), 293–317. https://doi.org/10.3102/0162373715616249

The Danielson Group. (2011). The framework for teaching evaluation instrument.