Teacher Education, Educational Leadership & Policy ETDs

Fall 12-3-2019

Teachers’ Perspectives of the Observation Process within the Evaluation

Christine E. Lenhart

University of New Mexico - Main Campus

Follow this and additional works at: https://digitalrepository.unm.edu/educ_teelp_etds

Part of the Educational Administration and Supervision Commons, Educational Assessment, Evaluation, and Research Commons, Educational Leadership Commons, and the Teacher Education and Professional Development Commons

Recommended Citation


This Dissertation is brought to you for free and open access by the Education ETDs at UNM Digital Repository. It has been accepted for inclusion in Teacher Education, Educational Leadership & Policy ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact amywinter@unm.edu, lслоane@salud.unm.edu, sarahrk@unm.edu.
Christine E. Lenhart
Candidate

College of Education
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Marjori Krebs, Chairperson

Kersti Tyson

Cheryl Torrez

Sheri Williams
Teachers’ Perspectives of the Observation Process within the Evaluation

By

Christine E. Lenhart

B.B.A., Finance, University of Notre Dame, 1998
M.Ed., University of Notre Dame, 2000
Administration Credential, University of Colorado, 2005

Dissertation

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
Multicultural Teacher and Childhood Education

The University of New Mexico
Albuquerque, New Mexico

December, 2019
Dedication

I dedicate this dissertation to my husband, Tony, and my children, Andrew and Kaitlin. They supported me even though this process entailed great sacrifices. Thank you for the endless encouragement, patience, and confidence.

I also dedicate this dissertation to my parents, Ruth Ann and Philip. They have both encouraged me throughout my educational journey.
Acknowledgements

I first want to acknowledge my family for being so supportive while I completed this Ph.D. program. My husband, Tony, my children, Andrew and Kaitlin, and my parents, Ruth Ann and Philip, provided me with the time to engage in learning. They cheered me on when this program felt challenging and endless.

I want to thank Dr. Marjori Maddox Krebs, my adviser throughout this Ph.D. program. Dr. Krebs continued to encourage me throughout this long journey and provide me with amazing advice and feedback. I also want to thank the members of my dissertation committee. Their continuous guidance and suggestions helped me improve the quality of my research and dissertation.

This study would not have occurred without the efforts of the five participants. These teachers shared their experiences and time with me. I was greatly inspired by their commitment to their students and school communities.
TEACHERS’ PERSPECTIVES OF THE OBSERVATION PROCESS WITHIN THE EVALUATION

By

Christine E. Lenhart

B.B.A., Finance, University of Notre Dame, 1998
M.Ed., University of Notre Dame, 2000
Administration Credential, University of Colorado, 2005
Ph.D., Multicultural Teacher and Childhood Education, University of New Mexico, 2019

ABSTRACT

One of the largest challenges for administrators is to create and implement teacher evaluation systems that help improve teaching and student learning. As policymakers and administrators have worked throughout the United States to develop and implement new teacher evaluation systems, teacher involvement in developing those teacher evaluation systems has been limited. Specifically, the evaluation process and the tools used in a state in the southwestern part of the United States to evaluate teachers do not incorporate teachers’ perspectives, yet they directly impact teacher standing, ability to receive tenure status, and, in some cases, their salary.

The purpose of this research study was to examine teachers’ perspectives. The qualitative data collected and analyzed during my phenomenological study reflected teachers’ perspectives of the observation portion of the evaluation system. This qualitative study collected data from five teachers who work in an urban school district. The participants are all elementary and middle school teachers in the same school district, located in the American Southwest. All school rankings in the school ranking system (A-F) are
represented in the group of participants. During the study, all of the participants recognized the need for incorporating observations into the evaluation system. However, the participants believed that changes needed to occur in order to increase the effectiveness and accuracy of the process. The changes suggested by participants included having evaluators spend more time in the classrooms, more evaluators visiting each classroom, and pre- and post-conferences. These adjustments could help teachers build stronger relationships with their evaluators and increase overall accuracy of collected data, as well as help evaluators know the students better. If these changes occurred, the participants indicated that the process would be more meaningful and helpful as they aim to strengthen their practices, which would ultimately benefit student learning.

This research is relevant for administrators and others who are developing and implementing teacher evaluation systems, helping them understand the teachers’ perspectives of the observation portion of the evaluation system when reflecting on the current system and making possible adjustments in the future.
# Table of Contents

**List of Tables** .......................................................................................................................... xiii

**Chapter 1: Introduction** ......................................................................................................... 1  
  Background of Study ............................................................................................................. 1  
  Noticed trends ....................................................................................................................... 3  
  Research goal ....................................................................................................................... 4  
  Personal experience ............................................................................................................. 5  
  Description of the Study ...................................................................................................... 6  
  Problem statement .............................................................................................................. 6  
  Significance of the study ...................................................................................................... 7  
  Research Questions ............................................................................................................. 7  
  Central question ................................................................................................................... 8  
  Sub-questions ..................................................................................................................... 8  
  Purpose of the Study ............................................................................................................ 8  
  Theoretical Framework ......................................................................................................... 9  

**Chapter 2: Literature Review** .............................................................................................. 12  
  History of the Teacher Evaluation System ......................................................................... 12  
  Teacher Effectiveness ......................................................................................................... 21  
  Qualities of effective teachers ............................................................................................ 24  
  Evaluation and Assessment of Teachers and Teaching ..................................................... 26  
  Accountability, authenticity, and reflection ......................................................................... 26  
  Criteria of the evaluator ...................................................................................................... 29  
  Prioritizing observations .................................................................................................... 31
<table>
<thead>
<tr>
<th>Chapter 3: Methodology</th>
<th>52</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methodological Framework</td>
<td>52</td>
</tr>
<tr>
<td>Phenomenology</td>
<td>52</td>
</tr>
<tr>
<td>Epoché</td>
<td>53</td>
</tr>
<tr>
<td>Intentionality of consciousness and noema</td>
<td>54</td>
</tr>
<tr>
<td>Variations of phenomenology</td>
<td>55</td>
</tr>
<tr>
<td>Research Activities</td>
<td>57</td>
</tr>
<tr>
<td>Epoché and transcendental-phenomenological reduction</td>
<td>57</td>
</tr>
<tr>
<td>Understanding positionality</td>
<td>59</td>
</tr>
<tr>
<td>Limitations</td>
<td>59</td>
</tr>
<tr>
<td>Delimitations</td>
<td>61</td>
</tr>
<tr>
<td>Participant selection</td>
<td>63</td>
</tr>
<tr>
<td>Methods of data collection</td>
<td>68</td>
</tr>
<tr>
<td>Methods of analysis</td>
<td>70</td>
</tr>
<tr>
<td>Conclusion</td>
<td>73</td>
</tr>
</tbody>
</table>

| Chapter 4: Catie | 74 |
Overall Teaching Evaluation Experience ................................................................. 74
Observation Process within the Teacher Evaluation ................................................. 75
Purpose of the Teacher Evaluation ......................................................................... 78
Effectiveness of the Observation Instrument .......................................................... 78
Components of the Observation Process ................................................................. 80
Data Collected during the Observation Process ..................................................... 81
Changes to the Observation Process ....................................................................... 82
Final Thoughts ......................................................................................................... 84

Chapter 5: Debbie .................................................................................................... 85
Overall Teaching Evaluation Experience ................................................................. 85
Observation Process within the Teacher Evaluation ................................................. 86
Purpose of the Teacher Evaluation ......................................................................... 87
Effectiveness of the Observation Instrument .......................................................... 88
Components of the Observation Process ................................................................. 88
Data Collected during the Observation Process ..................................................... 89
Changes to the Observation Process ....................................................................... 90
Final Thoughts ......................................................................................................... 91

Chapter 6: Kaitlin ................................................................................................... 93
Overall Teacher Evaluation Experience ................................................................. 93
Observation Process within the Teacher Evaluation ................................................. 93
Purpose of the Teacher Evaluation ......................................................................... 94
Effectiveness of the Observation Instrument .......................................................... 95
Components of the Observation Process ................................................................. 95
Chapter 7: Mary Gene ................................................................. 100
  Overall Teaching Evaluation Experience ......................................... 100
  Observation Process within the Teacher Evaluation ....................... 100
  Purpose of the Teacher Evaluation ................................................. 103
  Effectiveness of the Observation Instrument .................................... 103
  Components of the Observation Process ....................................... 104
  Data Collected during the Observation Process ............................ 105
  Changes to the Observation Process ............................................. 106
  Final Thoughts ........................................................................... 107

Chapter 8: Melinda ........................................................................ 109
  Overall Teaching Evaluation Experience ......................................... 109
  Observation Process within the Teacher Evaluation ....................... 111
  Purpose of the Teacher Evaluation ................................................. 111
  Effectiveness of the Observation Instrument .................................... 113
  Components of the Observation Process ....................................... 113
  Data Collected during the Observation Process ............................ 115
  Changes to the Observation Process ............................................. 115
  Final Thoughts ........................................................................... 116

Chapter 9: Findings ...................................................................... 119
  Theme 1: The Relationship between the Teacher and Evaluator ....... 119
Chapter 10: Discussion

Participants Perceptions of the Impact of Their Relationships with Evaluators

Participants Express Frustration and Anger When Talking About Their Observations

Participants Question the Accuracy of Classroom Visits

Evaluators Need to Know the Students

Conclusion

Chapter 11: Recommendations

Recommendations for School Administrators and School District Personnel

Recommendations for Teachers

Recommendations for Future Research

Summary

Appendices

Appendix A: Interview Questions

Appendix B: Observation Rubric

Appendix C: Domains

References
List of Tables

Table 1 Sample of Codes and Sub-codes .......................................................................................... 72
Chapter 1: Introduction

Background of Study

I still remember leading a faculty meeting in a public school in Northern California focused on the teacher evaluation system during my first year as an administrator over 15 years ago. As I began to embark upon the topic, some teachers immediately put their hands in the air while others interrupted me. I had led many meetings during the year and never felt this type of anxiety and frustration surface as I introduced a topic. I stopped my presentation and allowed teachers to share their thoughts.

Teachers discussed the need for a system that incorporated their personal professional goals. Most of them did not think the current evaluation system helped them grow. They felt it was simply an additional task they had to fulfill. Others questioned the ability of an evaluator to understand their classrooms after just a couple of visits. In summary, the teachers felt that evaluators needed to be experts in the teachers’ content areas, they needed to visit the classrooms more often, and the evaluation needed to focus on each teacher’s specific goals. They did not think the current evaluation helped them become more effective teachers.

Instead of continuing with my initial presentation, initiated by the district office, I began making a web diagram on the front board. The teachers added their ideas about how to improve the evaluation system. By the end of this activity, a crazy web diagram encapsulated the entire board. Teachers’ ideas clearly differed from the original ideas given by the district office. Throughout the day, I continued to visit that board and ponder different ways to include the ideas of the teachers in the process. As I looked at the board, I also flipped through the original PowerPoint presentation offered by the school district, actually
giggling to myself. There was very little overlap between the teachers’ ideas and those of the
district office. I knew, at this moment, that something had to change.

I returned to my office at the end of the day and called the superintendent. He asked
what the teachers thought of the teacher evaluation presentation. I replied that teachers
interrupted me as soon as I embarked upon the topic, and that they had expressed
immediate frustration. I continued to tell him about the activity and conversation. At first,
he was not open to the idea of collecting teacher feedback or changing the process. I asked
him to consider a one-year teacher evaluation pilot project at my school. I mentioned that we
could have half of the teachers continue on the same evaluation and half with a new
evaluation that incorporated their ideas. After many discussions, he agreed.

The other building administrator led the original evaluations, and I developed a new
evaluation with the help of the teachers. This evaluation incorporated both district goals and
the teachers’ personal goals. I visited each classroom at least once every other week.
Evaluators need to visit classrooms often to understand the classroom environment (Mendels,
2012). I provided informal feedback after every observation. I also met with teachers to
discuss the observations multiple times throughout the school year. These discussions led
teachers to adjust their goals, seek additional help from experts, and engage in video
exercises. Teachers involved also watched each other teach and provided feedback. That
year, we all worked hard to create a community that helped each other become better
educators.

At the end of the year, I provided all the teachers involved with feedback forms.
Teacher feedback indicated that those teachers who participated in the new evaluation system
appreciated its components and felt that they grew as educators within that one year. Most of the teachers who did not participate asked for the opportunity to do so the following year.

I presented the findings to the district administrators, explaining the new system and showing the feedback gathered from teachers. The final evaluation ended up being a compilation of all the observations throughout the school year, coupled with the district goals and personal goals of the teachers. Other administrators asked to implement the new evaluation, and those requests were granted by the superintendent.

Multiple teachers tried the new evaluation the following year and, again, the feedback from the teachers, that following year, showed their appreciation for an evaluation that truly represented their abilities and goals. I left the district after those two years, but this experience made me realize the power of asking for teachers’ perspectives on the evaluation system. When the system reflected teachers’ ideas and goals, they took ownership in the system and engaged in learning through the tool.

**Noticed trends.** As I continued to serve as an administrator, this situation continuously replayed in my mind and impacted my future experiences. When I began my Ph.D. program, I knew that I wanted to write my dissertation on the teacher evaluation process. I researched the history of teacher evaluation systems utilized throughout the United States, different strategies incorporated in various evaluation systems, and current policies and changes that impacted these evaluation systems. I dug even deeper to find pilot studies and different research projects that incorporated teachers’ perspectives. I wanted to compare their findings with the findings of my two-year pilot study in Northern California. I wondered if, by incorporating the teachers’ perspectives, other districts found that teachers developed ownership in the process and overall success.
As I continued to search, I was shocked that teachers’ perspectives were not incorporated in past or present research. I could not find documents that summarized teachers’ ideas about evaluation systems. I found articles about policymakers’ ideas, yet nothing about how teachers’ perspectives were incorporated into the development of the evaluation tools.

**Research goal.** Teachers engage in various evaluation processes throughout the United States every few years. Research shows that evaluators struggle to accurately define, monitor, and assess the success of teachers, often omitting areas of needed improvement (Marshall, 2005, Stake, 1980). Typically, the evaluation process involves only a few informal observations, which lead evaluators to rely upon limited data to assess teacher performance, making their evaluations of little use for teachers (Donaldson, 2010). Pitler and Goodwin (2008) asserted that school principals “understand the power of sharing their aggregated observation data with school staff to support professional learning communities” (p. 11). However, often times, school principals struggle to find time to collaborate with teachers and share their observations. When observing classrooms and noticing different strategies, the evaluators need to ask teachers questions to better understand the intentions behind the use of each strategy (Donaldson, 2010). After extensively researching teacher evaluation systems throughout the United States, I struggled to find research that incorporated teacher perspectives. As states continue to face challenges with the teacher evaluation process, I focused my research on better understanding teachers’ perspectives about the current challenges and successes within the evaluation process, and ways to improve the process.
My study filled the gap in the research, covering teachers’ perspectives on the observation portion of the teacher evaluation system. My research was focused on discovering teachers’ perspectives of the observation process, as part of the teacher evaluation system, in an urban school district in the Southwest.

**Personal experience.** As my introductory anecdote illustrates, I came to understand the need for this type of study, not only from the past and present research, but also from my personal experiences as a teacher and administrator. After teaching in six different states, I realized that various evaluation systems are used throughout the country. In my experience, some of the systems solely consist of a one-page checklist. Other evaluation systems incorporated a short discussion with the evaluator and, possibly, a brief written summary. During my evaluation experiences, only once did I interact with an administrator who had experience teaching at my grade level. Normally, the evaluators understood very little about the challenges that I encountered with my grade level and content areas. During my eight years as a teacher, I never had an evaluator visit my classroom more than twice in one year. After spending a brief time watching me teach, I never understood how an evaluator truly grasped what was going on in my classroom. Every year, school principals told me that I did a great job, but never provided feedback to help me improve. I needed and wanted to improve, however the evaluation systems did not help me reflect on my practices. I believed that the evaluators just wanted to get the evaluations completed, so they could move on to different administrative tasks. Because I wanted to improve, I was frustrated that the evaluative tools were not used to encourage ongoing dialogue or develop professional goals.

This experience and others led me to train as an administrator. I wanted to remember these feelings I felt as a teacher and act mindfully to advocate for the teachers while serving
as an administrator. As indicated in my introduction, I worked hard my first few years to listen to the teachers’ perspectives about the evaluation system and see how integrating their ideas increased overall ownership and results. Anonymous surveys, given to the teachers after the evaluation cycle, indicated that they truly appreciated the evaluative tools used as part of the pilot project over the two-year period. When I moved to other schools, I always asked questions about the evaluation systems. During my first two years as an administrator, a superintendent in Northern California was open to ideas because he received flexible guidelines from the state.

In summary, my personal frustration with receiving limited feedback from evaluators when teaching, and my experience as an administrator with making adjustments to the evaluation system to reflect teacher needs, led me to study the teacher evaluation system in more depth. I studied teachers’ perspectives about the classroom observation process within the teacher evaluation process in order to better understand their ideas about the successes and challenges of the current system.

Description of the Study

Problem statement. In my experience, teachers participate in the evaluation process either every year or every few years. Evaluators complete the forms provided by the school district offices. The forms vary from state to state. The development of the teacher evaluation process incorporates the ideas of policymakers and others not directly involved in education (Danielson, 1996; McCaffery, Lockwood, Koretz, & Hamilton, 2003; Stake 1980). As evaluation tools focus on assessing teacher performance, only a limited number of teachers receive the opportunity to state their ideas on the development or implementation of those tools (Danielson, 2001). Currently, teachers have not been asked about their
perspectives of the observation component of the current evaluation system and how it impacts their teaching.

**Significance of the study.** This dissertation project focused on learning teachers’ perspectives about the challenges and successes involved in the observation process within the evaluation system used in an urban school district in the Southwest. In my opinion, teachers’ perspectives are important because teachers play the central role in the evaluation process. In order for the tool to effectively impact student learning, teachers need to value the tool and use the feedback to create goals and learning opportunities (Danielson, 1996; Knapp, Honig, Plecki, Portin & Coplan, 2014). By listening to the perspective of teachers, the evaluation process would better align with their professional goals and become an active document (Heneman & Milanowski, 2001). Multiple measures must be incorporated during the evaluation process in order to accurately assess teachers’ practices and the impact on student performance (Stiggins & Bridgeford, 1985). Evaluators must be highly trained to assess performance and help the teachers develop authentic professional goals. In order to encourage teacher professional growth, a successful evaluation process is essential (Darling-Hammond & Snyder, 2000; Orr, Berg, Shore, & Meier 2008). In summary, the purpose of this study was to understand how teachers in an urban school district in the Southwest perceive the observation process within the teacher evaluation system, to discover the challenges and successes through their perspectives, and to determine the ways in which teachers hope to see the evaluation system change in the future.

**Research Questions**

My research questions were divided into a central research question and two sub-questions.
Central question. My central research question was: *What are teachers’ perspectives on the observation portion of the teacher evaluation process in a southwestern, urban school district?* The sub-questions and interview questions all stemmed from this main question. The central question provided the primary focus for this study.

Sub-questions. My sub-questions included the following:

1. From the teachers’ perspectives, how effective is the observation portion of the teacher evaluation system?

2. What areas of the observation portion of the evaluation system are most useful from the teachers’ perspectives?

From these sub-questions, ten interview questions (see Appendix A) were developed to ask study participants.

Purpose of the Study

The purpose of this phenomenological study was to describe the experiences and perspectives of teachers involved in this urban, southwestern school district’s evaluation process. More specifically, I learned about their overall experiences with the evaluation system and the challenges and successes of the observation process.

This study matters because teachers’ perspectives currently have limited influence on the development and implementation of the evaluation system, yet evaluators expect teachers to participate in the system to grow professionally. When teachers engage in evaluation processes and grow professionally through collaboration and reflection in their teaching, student academic performance is enhanced (Danielson, 2001).
Theoretical Framework

Theoretical frameworks illustrate an approach that directs the researcher throughout the study (Corbin & Strauss, 2008). Social constructivism was the theoretical framework that guided my qualitative research project. The social constructivist perspective focuses on using interactions and dialogue with others to create knowledge. As I focused on learning about teacher perspectives, I centered my research on engaging in dialogue with teachers to better understand their experiences with the observation process. I believed that teachers need collaboration and multiple interactions with evaluators, other teachers, and students to develop their understanding while incorporating their own background, knowledge, and culture. Through the social interactions, teachers continue to formulate their understanding of best practices (Bruner, 1990; Richardson, 1997).

Social constructivist perspective influenced my methods throughout this dissertation. By interviewing teachers, I learned more about teachers’ perspectives of the evaluation process and how their perspectives impact their own learning and the learning of their students (Merriam 2009, Moustakas, 1994). The social constructivist perspective guided my interview preparation and long interviews with teachers. By using open-ended questions and being open to their responses, I focused on learning about the way teachers view the observation process through their own lenses (Creswell, 2013).

Dewey (1938) believed that education is a social and interactive process and should focus on connecting knowledge to personal experiences and interests. He felt the primary goal in education was to prepare people to thrive in their community and be productive members of society. He mentioned that, in order for people to thrive, they need to learn how to think independently, critically, and collectively with their community.
Social constructivist theory informed my perspective of the teacher evaluation process. Interaction and collaboration must be incorporated into the evaluation system for teachers to grow (Dewey, 1938). Additional collaboration among disciplines enhances the learning process (Palincsar, 1998). Teachers rarely interact with the evaluators, peers, or students to discuss their performance during the evaluation process. Normally, the evaluator visits the classroom twice during the school year, and then the evaluator tells the teacher the findings during the observations, at the end of the school year. During the conversation about findings, little collaboration between the teacher and evaluator takes place (Danielson, 1996). Dewey’s (1938) research suggested that by increasing interaction and collaboration among teachers and evaluators, teachers will reflect on their practice and grow professionally. When evaluators engage with teachers in collaborative exercises, both parties learn from the discussed experiences (Danielson, 1996; Darling-Hammond, 2014; Dewey, 1938). The lack of collaboration within the evaluation system leads teachers to frequently dismiss the feedback (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Stake, 1980; Darling-Hammond, Wise, McLaughlin, & Bernstein, 1985). When a teacher does not receive an opportunity to explain and/or discuss the feedback, the feedback holds little value (Frase, & Streshly, 1994; Stake, 1980). Collaborative and reflective opportunities during the evaluation process help teachers learn from prior experiences and construct new knowledge (Gallagher, 2004; Krechevsky, Rivard, & Burton, 2010).

Summary

My personal experiences, serving as teacher and administrator in private and public schools, and my understanding of the current research on the teacher evaluation process both led me to study teachers’ perspectives of the observation component of the teacher evaluation
process. This study added to the body of knowledge because it provided information about the teachers’ perspectives on one specific component of the evaluation process. Teacher perspectives provide first-hand insights about the successes and challenges of the observation tools. The purpose of this study was to gain insight into the ways teachers view the current observation process.

In this dissertation, chapter two focuses on the literature review, while chapter three discusses the methodology. Chapters four through eight encompass the individual input gained from each study participant. Each participant is honored by having their own chapter to illustrate the individuality of their experiences. Chapter nine reviews the findings and chapter ten focuses on the discussion. Lastly, chapter eleven discusses the recommendations.
In order for the classroom observations to effectively impact student learning, teachers need to value the tool and use the feedback to create goals and learning opportunities (Danielson, 1996; Danielson, & McGreal, 2000). By listening to the perspective of teachers, the evaluation process can more clearly align with their professional goals and becomes an active document (Heneman & Milanowski, 2001). Multiple measures must be incorporated during the evaluation process in order to accurately assess teachers’ practices and the impact on student performance (Bill & Melinda Gates Foundation, 2013; Stiggins & Bridgeford, 1985). “Teaching is too complex for any single measure of performance to capture it accurately” (Resmovits, 2013, p. 10). Evaluators must be highly trained to assess performance and help the teachers develop authentic professional goals (Darling-Hammond, 2000). Currently, many districts throughout the United States are requiring evaluators to complete training prior to evaluating teachers (Leahy, 2012). “Policies that create increasingly valid measures of teacher effectiveness— and that create innovative systems for recognizing, developing, and utilizing expert teachers—can ultimately help to create a more effective teaching profession” (Darling-Hammond, 2013, p. 171). This literature review focuses on the history of supervision and teacher evaluation, teacher effectiveness, and different types of evaluations and assessments.

History of the Teacher Evaluation System

The development of supervision and teacher evaluation systems started in the 1700s. In the 1700s, people used current town structures and appointed various clergy members or government officials to act as supervisors (Tracy, 1995). These supervisors hired teachers for schools, built the curriculum they saw fit for the community, and evaluated teacher
performance. Most of the feedback focused on teachers’ ethics, rather than instruction. Due to the lack of expertise in pedagogy and instruction, the supervisors provided a large range of feedback (Shaw, 2016).

As the education system became more complex in the mid-1800s, there developed a need for more specialized teacher supervisors. Large urban areas formed due to the rise in industrial development throughout the United States. With the development of industrial areas and the common school movement, large school districts formed, needing more teachers and supervisors with expertise (Tracy, 1995). This need led to the role of principals, who were educated in pedagogy, understood the challenges found in various classrooms, and served as school leaders. In both urban and rural communities, educators began to hold more specific roles at schools, helping schools to run more efficiently. Educators quickly realized the need for more specific feedback as the roles of teacher and supervisor became more complex, requiring expertise in various subject areas and teaching skills. Supervisors started focusing on providing feedback to help improve and enhance teaching strategies (Blumberg, 1985). In order for educators to serve as experts in the field, teachers and supervisors needed to acknowledge the need for pedagogical skills (Marzano, Frontier & Livingston, 2011).

Early in the 1900s, two primary views guided the discussions on education. John Dewey (1938) viewed democracy as the cornerstone for human development and progress. In schools, he advocated that students focus on understanding democracy by connecting their in-school realities to real-world experiences. He believed that students’ roles in school should mirror those that they would later fill in society. The second view focused on scientific management. Fredrick Taylor (1911) analyzed the work habits of factory workers and how their behaviors impacted overall productivity. He examined parallels between the
factories and schools. He believed that even though there were numerous ways tasks could
be fulfilled, a best method could still surface (Taylor, 1911).

In 1918, based on the ideas shared by Taylor, John Franklin Bobbitt wrote a book
called *The Curriculum*. The book, which outlined standards for designing and building
curriculum, was widely used throughout the United States. “To know what to do is as
important as to know how to do it” (Bobbitt, 1918, pg. v). Similar to Taylor, Bobbitt felt that
an effective school should mimic the structure of a factory. The principal should manage the
organization, the teachers should serve as supervisors, and the students should act as workers.
Teacher actions and behaviors operated as the input, while student academic results served as
the output. Educators began to use measurements of inputs and outputs as a scientific
approach toward improving school organizations and educational practices (Cubberley,
1929). Scientific management “focused on judging quality and efficiency through a focus on
inputs and outputs” (Shaw, 2016). Also using the foundation of Taylor’s factory structure,
Cubberley developed criteria for supervisors to follow during classroom observations, using
a scale from A to F. Supervisors assigned a letter to the lesson and provided specific
feedback, focused on weak areas and suggestions for future lessons.

In the 1950s, many educators supported the Tylerian model, which focused on
curriculum planning. This framework focused on developing an objective, selecting
intentional activities to fulfill that objective, implementing the activities effectively, and then
assessing the implementation (Tyler, 1950). From this model, the first checklist used during
classroom observations was developed. The checklist focused on specific teacher behaviors
and judged overall teacher effectiveness (Marzano, Frontier, & Livingston, 2011). Many
educators began to debate how much teacher effectiveness and curriculum planning directly impacted student learning.

Clinical supervision developed in the 1960s. Goldhammer (1969) and Cogan (1973) led the popular discussion during this decade. Robert Goldhammer (1969) developed a five-phase process focused on developing a reflective dialogue between teachers and supervisors. The phases included pre-conferences, classroom observations, analysis of the data, post-conferences, and then analyzing the analysis. This five-phase process focused on the relationship of the teacher and student and how the relationship made an impact on overall student learning. The phases promoted collaboration between teachers about effective practices. Goldhammer’s professor, Morris Cogan, expanded the research even further. He discussed classroom behaviors and the need for supervisors to look for specific learning situations that led students to engage in critical learning. He believed that the supervisor’s relationship with teachers should focus on teacher learning and overall growth (Cogan, 1973). Cogan also mentioned how supervisors observe classrooms with a bias towards their own models of teaching, which could negatively impact their ability to provide accurate feedback. “Most teachers have consciously and unconsciously constructed a personal model of the good teacher. Such conceptions generally grow by accretion rather than by critical examination and careful testing” (Cogan, 1973, p. 54). Neither Cogan nor Goldhammer structured their observations prior to classrooms visits, instead developing categories and the needed structure during and after the observation. This allowed them to avoid preconceived ideas of what the structure should look like before observing the behaviors in the classroom (Cogan, 1973, Goldhammer, 1969).
In the 1980s, Madeline Hunter developed a behavioral seven-step lesson plan, also called the Direct Instruction Model, which included having pre- and post-conferences with the supervisor. The seven steps included: anticipatory set, objective and purpose, input, modeling, checking for understanding, guided practice, and independent practice (Hunter, 1982). Hunter believed that the conferences should focus on discussing classroom behaviors, instructional practices, continued growth, and data supporting the behaviors. This model also included script taping. Supervisors used script taping to write down observed behaviors, and then coded and placed them in categories to determine if the teacher used effective or ineffective techniques to promote student learning. The Hunter Model helped teachers plan instruction and supervisors lead evaluations. Hunter also discussed the importance of observing teachers differently, dependent on their levels of experience. She believed in differentiating supervisory techniques, as well.

In 1965, the Elementary and Secondary Education Act (ESEA) was enacted in the United States with the primary objective of providing greater consistency and uniformity across the country’s schools, no matter the location or associated income demographics. This policy stemmed from President Johnson’s War on Poverty. In 1983, the report *A Nation at Risk* focused on failing schools in the United States and how they were inferior to schools in other countries around the world (National Commission on Excellence in Education, 1983). The educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people. What was unimaginable a generation ago has begun to occur --- others are matching and surpassing our educational attainments (National Commission on Excellence in Education, 1983, p. 5).
The report focused on increased accountability, mentioning the need for higher standards and specifically the implementation of teacher evaluation processes to monitor teacher quality. The report discussed how teacher salaries, tenure, and retention should be tied to evaluations, and recommended that all teachers take basic skills and content-specific assessments. Poor evaluations, the report asserted, should lead to termination in order to increase teacher quality. Following this report, some states mandated observation systems. This was an attempt to increase teacher effectiveness in classrooms across the country, though it remained unclear how effective the programs were at preparing or improving the work of teachers (Education Commission of the States, 2014).

In response to the clinical supervision models, in the mid-1980s, William Glatthorm (1984) discussed the importance of teacher input in the evaluation system. Glatthorm asserted that teachers needed to develop career goals and take ownership in their professional development. He also discussed the importance of differentiating the supervisory system for each teacher.

Thomas McGreal (1983) expanded on these ideas. He believed that ineffective teachers and new teachers have different needs than effective teachers. Intensive programs were developed for high-stake decision-making processes, focused on tenure and continued employment. Ultimately, the primary goal of supervision was to improve overall instruction, to positively impact student learning (Glickman, 1985). In order to improve instruction, Glickman discussed the importance of providing teachers with needed assistance, professional development opportunities, and applicable action research. During the 1980’s, research focused primarily on the need for an effective teacher evaluation system and argued against clinical supervision.
In 1996, a study indicated that a majority of the teacher evaluation systems across the United States were being utilized for teachers who were struggling or at risk of being dismissed. Evaluations were not focused on helping the overall population of teachers to develop and grow. The study also indicated that very little research was used in the development of teacher evaluation processes (Loup, Garland, Ellett & Rugutt, 1996). RAND also completed a study focused on the types of evaluation systems used in schools across the country. One finding from this study showed the need for specific, detailed feedback to help teachers grow. “In their view, narrative evaluation provided insufficient information about the standards and criteria against which teachers were evaluated, and resulted in inconsistent ratings among schools” (Wise et al., 1984, p. 71). Other concerns, identified during the study, included a lack of supervisory expertise and training, the unwillingness of teachers to embrace provided feedback, and the inconsistency of the evaluation processes.

In the 1990s, the focus shifted from analyzing solely the teacher to the broader view of the relationships developed between the teacher and students. Danielson (1996) introduced the Framework for Teaching, based on prior work with preservice teachers, which included observation rubrics. She focused on learning about the competence of these teachers and various classroom dynamics. Danielson’s rubrics included four domains that defined effective teaching within a framework of common language for educators to use when discussing various educational topics. This framework helped teachers self-assess and reflect on their practices (Marzano, Frontier, & Livingston, 2011). The rubric included multiple components of teaching and could be applied to all grade levels and disciplines. Each component was measured using four different grading levels.
In 2001, the United States Congress passed the No Child Left Behind Act (NCLB) and the primary focus became accountability measures and evaluation, rather than supervision. The development of NCLB came from research stating that student growth is most greatly impacted by teacher quality and effectiveness (Kane & Staiger, 2008; Rockoff, 2004). In order to measure teacher quality, Congress shifted from looking at teachers’ years of experience and degrees attainment to using student test scores (The Education Commission of the States, 2014). The importance of student achievement became a large component of the teacher evaluation system (Tucker & Strong, 2005). “Given the clear and undeniable link that exists between teacher effectiveness and student learning, we support the use of student achievement information in teacher assessment” (Tucker & Stronge, 2005, p. 102). By 2014, as stated in the NCLB, all students were expected to achieve proficiency. To help achieve this goal, NCLB developed a waiver process and grant competition. The grant competition was called Race to the Top (RTTT). NCLB required states to change their teacher evaluation systems, increasing the accountability of individual teachers by focusing on data. Tennessee was one of the first states to adopt the value-added models. In Tennessee, 25-35% of the evaluation included measuring teacher performance using value-added scores (The Education Commission of the States, 2014).

In 2009, a nonprofit organization called the New Teacher Project shared a report focusing on 14 school districts. The report stated that 99% of teachers received a satisfactory rating on their evaluations (Weisberg, Sexton, Mulhern, & Keeling, 2009). “When all teachers are rated good or great, those who are truly exceptional cannot be formally identified” (Weisberg et al., 2009, p. 6). At that point, a surplus of money for additional research was filtering into the education system from the Bill & Melinda Gates Foundation,
which greatly influenced policy (Shaw, 2016). The foundation also developed the Measures of Effective Teaching (MET) project which researched effective tools and trainings used to observe teacher quality (Bill & Melinda Gates Foundation, 2013).

In 2009, a study associated with the New Teacher Project, “The Widget Effect,” critiqued the teacher evaluation system across the country. “The Widget Effect describes the tendency of school districts to assume classroom effectiveness is the same from teacher to teacher” (Weisberg et. al., 2009, p. 4). This study was conducted in 12 districts across four different states. The findings of the study illustrated that teachers do not feel the evaluation process caters to their professional developmental needs nor does it provide useful information to support growth. The conclusion discussed the need to rebuild the evaluation system. (Weisberg et al., 2009).

In 2014, the National Council on Teacher Quality (NCTQ) analyzed curriculum used in 836 teacher preparation institutes, stating that strong preparation programs clearly link to teacher effectiveness and positive change in the education system. However, to develop effective preparation systems, educators must define effectiveness, develop useful data systems, and share data findings across states (The Education Commission of the States, 2014). “According to DQC [Data Quality Campaign], 46 states say they can now create reports that include longitudinal statistics on school systems and groups of students to guide school-, district-, and state-level improvement efforts” (The Education Commission of the States, 2014, p. 4). Because some state laws do not allow sharing of data, there are intrinsic challenges with this program and its success (The Education Commission of the States, 2014).
In 2015, President Obama signed a new act called the Every Student Succeeds Act (ESSA). NCLB and ESSA are similar in some ways, but ESSA “does not require states to set up teacher-evaluation systems based in significant part on students’ tests scores” (Sawchuk, 2016, p. 1). The ESSA focuses on closing achievement gaps, providing equitable education of all students, and preparing students for college and careers. This new act, ESSA, provided states with opportunities to revise their current grading systems to assess teacher performance in new ways.

In summary, as the history of supervision and teacher evaluation started in the 1700s, it continues to change, evolving over the years. Throughout decades, the primary focus has moved among many areas, including general teacher feedback, curriculum planning, observation with pre- and post-conferences, overall lesson planning, and relationships between teachers and students. In recent years, the focus of teacher effectiveness and accountability has moved increasingly toward student achievement. It is clear that as educational processes change, so will supervision and teacher evaluation strategies.

**Teacher Effectiveness**

Effective teaching must first be defined before it can be accurately assessed (Danielson, 1996). In order to understand how to build and implement an evaluation process and measure effectiveness to enhance student learning, components of effective teaching must first be determined. Danielson (1996), Marzano, Pickening and Pollock (2001), Varlas (2009), and Goe, Bell and Little (2008) discussed different aspects and rubrics used to understand and analyze teacher effectiveness. Danielson examined teacher effectiveness and ways to enhance professional practices. Danielson developed a framework with four domains to demonstrate “what teachers should know and be able to do in the exercise of their
profession” to illustrate effective teaching (1996, p. 4). Danielson’s framework breaks down different aspects of teaching, which helps educators and policymakers recognize and share common vocabulary when analyzing effectiveness. The four domains include planning and preparation, classroom environment, instruction, and professional responsibilities (Danielson, 1996; Danielson, 2013). In order to achieve success within planning and preparation, teachers must become experts in their content areas, understand characteristics and learning needs of their students, select appropriate goals, design effective instruction, and assess student progress. Within the classroom environment, all members of the community must act in a respectful manner and maintain a culture of learning. Teachers must develop classroom procedures and manage behaviors appropriately. The physical space must be safe and arranged properly for all resources to be accessible to students. Teacher instruction needs to be clearly communicated, while providing engaging questions and content. Students need to frequently receive feedback from teachers and peers. Lastly, teachers need to communicate with families, continue to grow, and act professionally (Danielson, 2013).

In Classroom Instruction that Works, meta-analysis, which is a collection of results from a number of studies that explain the average effect of the studied strategy, helps identify strategies with the highest probability of success when analyzing student achievement (Marzano, Pickering & Pollock, 2001). For example, students engaged in higher-level questioning scored 0.73 standard deviations above those students not involved in higher-level questioning. Marzano, Pickering and Pollock (2001) discussed nine research-based strategies that improved student achievement, including: identifying similarities and differences, summarizing and notetaking, reinforcing effort, providing recognition, homework and practice, nonlinguistic representations, cooperative learning, setting objectives, and providing
feedback, generating and testing hypotheses, and questions and advanced organizers. With Marzano’s framework, evaluators look to see if teachers integrate the educational strategies intentionally (Pitler & Goodwin, 2008). Marzano, Pickering, and Pollock use these nine strategies to define aspects of effective teaching and their impact on student achievement.

While completing research for the National Comprehensive Center for Teacher Quality (NCCTQ), Goe, Bell, and Little (2008) defined five aspects of effective teachers: a) Effective teachers do not only focus on student academic success; they also focus on helping students succeed in all aspects of school, including social situations and daily attendance; b) Effective teachers have high expectations and help all students achieve the expectations; c) Effective teachers create intentional learning opportunities for students that engage students, assess understanding, and differentiate instruction; d) Effective teachers create diverse and civic-focused class cultures; and e) Effective teachers work well with other adults in the school community and collaborate with them about student needs.

“Although the research does not indicate that the passage of time will make all teachers better or incompetent teachers effective, it does indicate that, for most teachers, effectiveness increases with experience” (Varlas, 2009, p.1). Selecting teachers intentionally and carefully, as well as training them, will help improve teacher effectiveness when entering the teaching field.

If students work with highly effective teachers for three or more consecutive school years, their standardized test scores will be about 50 percentile points higher than the students who worked with less effective teachers (Sanders & Rivers, 1996). As a result, defining aspects of an effective teacher becomes pinnacle for helping increase student
achievement. In order to develop a tool to evaluate teachers, educators must first define and understand components of effective teaching (Varlas, 2009).

**Qualities of effective teachers.** Frameworks provide common vocabulary for discussing and understanding teacher effectiveness. Various teacher qualities are analyzed to better understand teacher effectiveness and the impact on overall student achievement (Gallagher, 2004; Kini & Podolsky, 2016; Rice, 2003). Rice (2013) explained that teacher experience, preparation programs, teacher certification, coursework, and high test scores positively impact teacher effectiveness. The educational institution attended by the teacher, pedagogical coursework taken at that institution, and advanced degrees earned in teacher education also achieved positive effects. No positive effect was found between student achievement and emergency or alternative credentials. Rice identified that the complexities of teaching and the many factors that can influence student progress signify the need for a multifaceted teacher evaluation system that incorporates multiple measurements.

Kini and Podolsky (2016) analyzed 30 studies, with data collected data over 15 years, on the impact of teacher experience on teacher effectiveness. The research methods used to analyze the data included longitude data as well as student and teacher fixed effects. When using teacher fixed effects, the researchers compared a new teacher (in their first or early years of their profession) with that same teacher after they accrued more years of experience. They studied how the same teacher impacted student achievement differently throughout their career. “Gains in teacher effectiveness associated with (teacher) experience are most steep in teachers’ initial years, but continue to be significant as teachers reach the second, and often third, decades of their careers” (p. 1). Kini and Podolsky found that there is a positive association between teacher experience and student achievement. Also, as shown
when following standardized test scores, students learn more when they work with a more experienced teacher, and improve overall in regard to other aspects of school, including daily attendance.

Gallagher’s (2004) research focused on teachers’ skills and how those skills impacted student achievement. Gallagher analyzed the relationships between evaluation scores and student progress. The composite measure that assessed overall student progress included teachers’ scores in math, reading, and literature. During the first stage of the study, Gallagher found a statistically significant relationship between the teachers’ scores and student achievement in reading, and a strong relationship, but not statistically significant, between the teachers’ scores and student achievement in math. Lastly, Gallagher also discovered, through his research, that other teacher qualities, including licensure and years of experience, were insignificant predictors of the variation.

The empirical evidence is inconsistent in regards to the effects of teacher education and experience on student achievement (Wayne & Young, 2003). Wayne and Young completed a narrative synthesis of 21 studies. They concluded that, in history and language, there was not sufficient empirical evidence illustrating that teacher education directly impacts student achievement. In science, there was some empirical evidence, yet more research would be needed to draw a conclusion. In math, Wayne and Young concluded that the achievement of high school level students is impacted by a teacher’s education. The 21 individual studies did not focus specifically on math in elementary and middle school classrooms, so no conclusions were drawn. When looking at other research focused on teacher education and student achievement, the overall findings are mixed (Goldhaber, 2004).
Effective teachers positively impact student achievement, but it is still unclear as to which specific teacher qualities enhance overall effectiveness. Empirical evidence is mixed when discovering which qualities increase teacher effectiveness (Sanders, Wright, & Horn, 1997).

Evaluation and Assessment of Teachers and Teaching

**Accountability, authenticity, and reflection.** One of the largest challenges in the teacher evaluation system is building an authentic assessment that accurately measures teacher effectiveness. Darling-Hammond and Snyder (2000) used a variety of case studies from teacher education programs to illustrate the impact of *authentic* assessments on teacher practice. *Authentic* refers to tying teacher goals and professional development to the evaluation process, thus helping teachers to reflect on their practices. The case studies demonstrate how *authentic* tools aid in the development of influential and practical strategies for fostering contextualized teaching. Teachers develop a better understanding about the relationships between theories and practices when using *authentic* assessments. According to Krechevsky, Rivard, and Burton (2010), by reflecting on their practice, teachers develop additional skills for managing the increasing challenges found in the education community. The use of *authentic* teacher assessments helps teachers better prepare for the challenging curriculum and diverse student bodies.

In the late 1990’s, due to the increased challenges faced by teachers to cater to more diverse learners, multiple teacher education programs started using more *authentic* assessments including cases, exhibitions, portfolios, and action research. “These tools allow the application of theoretical principles to problems in specific contexts while appropriately complicating efforts to draw generalizations about practice” (Darling-Hammond & Snyder,
By assisting teacher learning through the use of assessments, teachers developed additional skills that helped them teach increasingly challenging curriculum and more diverse communities of learners. Teachers developed a better understanding about the relationships between theories and practices when using *authentic* assessments. *Authentic* teacher assessments supply evidence of teacher learning and encourage a deeper understanding of practice. While Darling-Hammond and Snyder focused on the importance of using *authentic* assessments of teaching, Krechevsky et al. (2010) discussed using documentation to develop three types of accountability in school environments.

Krechevsky et al. (2010) examined three ways of being accountable and reflective using documentation: accountable to self, accountable to each other, and accountable to the larger community. Documentation enhances the learning process by examining, documenting, and interpreting the teaching process using multiple types of media. To study the three different realms of accountability, Krechevsky et al. investigated Wickliffe Progressive Community School, a “parent-choice” school populated with a diverse student body (2010, p. 65). Teachers were accountable to themselves when they reflected on their practices and acted as meticulous students of their own teaching. Accountability to each other included all members of the school community, taking ownership and responsibility in their personal learning and the learning of others. By using documentation, the school community developed a unified identity. Accountability to the larger community involved sharing documentation outside of the school community to educate the larger population about student learning. The use of documentation supported the three areas of accountability, which resulted in more intentional and insightful teaching, and provided proof of student learning.
In the spirit of transparency and school accountability, in 1999 Florida was the first state that embraced an A-F ranking system for schools. As approximately 16 other states have now adopted an A-F framework, others have already started looking at system modifications. The main objective of this accountability system was to push schools to improve overall performance, which included teacher effectiveness and student achievement. However, “there has been very little research conducted to determine the reliability or validity of using them for state accountability” (“A-F school ranking,” p. 1). In 2013, some research was conducted at the University of Oklahoma and Oklahoma State University. This research stressed the concern that the letter grades were based on limited data and that more data points should be included to create an accurate school ranking. The researchers indicated that the rankings “did not explain the how or why of low performance, and the system did nothing to build the capacity of schools or educators” (A-F school ranking, p. 1). By using only limited data, the rankings did not always represent the needs of various subgroups located within the schools nor show movements towards improvement. Also, using only five rankings among a large array of schools leads very different schools to achieve the same ranking. When looking at various criteria within a school, including student achievement, teacher performance, student attendance, and dropout rates, using one composite score is difficult when representing a large set of school criteria. These rankings also led teachers and administrators to gravitate towards working at schools with higher rankings. “But while a [school] grade is superficially clear and simple, it is not necessarily ‘about school performance,’ in the sense that the superficial understanding can reflect patently invalid representations of school quality” (Horn & Murray, 2015, p. i). In addition,
the authors stated that schools need to be graded using multiple criteria in order to accurately measure improvement and increase overall accountability (2015).

Due to the complexity of teaching and developing *authentic* and reflective assessments that enhance teacher and school accountability, many have researched and analyzed various systems and their abilities to measure overall effectiveness.

**Criteria of the evaluator.** Researchers have focused on different evaluator criteria and how various criterion impacts the overall evaluation system and the teacher experience. The collage of experiences, expertise, and trainings of the evaluator can impact the overall quality of the evaluation (Gallagher, 2004).

Many researchers have analyzed the importance of training evaluators. Training all evaluators to use the evaluation tools, providing precise and prompt feedback to teachers, and coaching effectively during the observation process are extremely important when developing an *authentic* evaluation system (Stiggins & Bridgeford, 1985). Multiple states require training before evaluators engage in the evaluation process (Leahy, 2012; TeachNM, 2015). Evaluator training helps strengthen the skills for collecting and providing feedback to teachers and pushes evaluators to better understand strategies used in classrooms. Evaluators also learn the importance of collaborating after the observation, specifically the discussions surrounding how, when, and why. By discussing the how, when, and why, evaluators coach teachers to develop a deeper understanding of their practices and in turn the teachers are better able to reflect on how to improve. Teachers need to intentionally integrate strategies into their practices, like collaborative activities. By being intentional, teachers develop effective practices. By training evaluators to accurately use the evaluation tools and effectively coach teachers, teachers begin reflecting on their intentions and better understand
why they use certain strategies. Also, evaluators must use the tools consistently when observing multiple teachers (Pitler, 2007).

Many have analyzed the impact of choosing evaluators with pedagogical content knowledge to enhance the teacher evaluation process. As the first part of Gallagher’s study focused on quantitatively testing the impact of the evaluation system and teacher characteristics against student achievement, the second portion of Gallagher’s study included qualitative research. Gallagher (2004) interviewed teachers and evaluators and analyzed data to discover the factors that added to the strength of the evaluation system and student achievement in reading and math. The relationship between the scores and student achievement in reading is higher than in math because most teachers and evaluators have acquired more knowledge in reading than math, which led to stronger assessments and alignment of standards. Most importantly, this specific study illustrated that the subject-specific evaluations, handled by evaluators who had attained pedagogical content knowledge, increased the validity of those evaluations.

The evaluators must understand the content area and the community of students in the classroom (Wright, Horn & Sanders, 1997). For example, teachers in classrooms with a number of English language learners (ELL) or students with disabilities (SWD) use different instructional strategies and have unique challenges. “Principals spend time in classrooms to evaluate instruction or, especially in the case of secondary schools where they can’t reasonably be expected to be experts in all academic disciplines, ensure that someone who is qualified does so” (Mendels, 2012, p. 56). Evaluators need to be trained to know if the instructional strategies being used with various subgroups are appropriate and effective.
Principals can’t be experts in every area, so districts should assure that someone with the appropriate expertise observes each teacher (Mendels, 2012).

**Prioritizing observations.** Evaluators need to prioritize their observations and coaching opportunities to provide feedback to teachers (Donaldson, 2010). In order to collect accurate observation data, multiple observations need to occur (Resmovits, 2013; Wright, Horn, & Sanders, 1997). During evaluations, evaluators tend to observe only 0.1% of the teachers’ yearly teaching, which leads the evaluation system to be ineffective and inefficient (Marshall, 2005). “Administrators also face urgent matters requiring their immediate attention every day, so putting off regular visits to classrooms seems justified, and they hope that visiting classes for evaluation purposes fulfills at least a minimum standard for instructional leadership” (Ginsberg & Murphy, 2002, p. 2). Because of the limited observations, evaluators receive an incomplete picture of teachers’ practices. Furthermore, during observations, evaluators tend to act like they know more than they do about instructional strategies, and teachers tend to build lessons that do not accurately represent their performance on most instructional days. Evaluations such as this rarely lead to instructional improvement and changes, and do not impact student learning. Evaluators are required to complete detailed descriptions of their observations to fulfill various district requirements, thus they have little time to walk around a classroom or to focus on student learning. Evaluators often portray more favorable pictures to avoid grievances, and teachers do not accurately share their experiences in the classroom in fear of seeming unprofessional. When evaluators provide more positive feedback than is deserved, average work continues (Marshall, 2005).
Evaluators need to prioritize ample time to observe teachers. Working with 15 urban schools, the Northern U.S. Charter Management Organization focused on informal observations and post-observation coaching sessions to fulfill evaluation requirements. Evaluators provided teachers with differentiated coaching opportunities every other week. Both the evaluator and teacher completed a summative evaluation packet and discussed their perspectives on the observations during the year. The evaluator compiled data and information from both inside and outside of the classroom to complete the summative evaluation. Evaluators worked with a small group of teachers each year. Overall, evaluators realized the importance of allowing more time for collaboration and observation and providing meaningful feedback to the teachers, which led teachers to make positive changes to their teaching (Donaldson, 2010).

Cincinnati’s Teacher Evaluation System (TES), which emerged in Cincinnati, Ohio, used Danielson’s framework. The completion and implementation of the evaluation tool resulted from an agreement between the Cincinnati Board of Education and Cincinnati Federation of Teachers. The new system clearly stated the evaluation criteria by focusing on Danielson’s framework, which included the four domains. Within the system, teachers applied to serve as evaluators. When evaluating an experienced teacher, the administrator observed the teacher once. The teacher evaluator, who had expertise and content knowledge, observed the teacher three times. The evaluators scheduled two of the observations and arrived unexpectedly during the remaining two visits. New teachers and low-performing teachers were evaluated on a different type of tool called the Peer Assistance and Evaluation Program. In this program, consulting teachers coached and collaborated with the new and low-performing teachers. By dividing the evaluation criteria among consulting teachers,
evaluating teachers, and administrators, administrators were able to provide higher quality data to teachers. Also, all members of the process received trainings every other week to analyze standards and had to pass a certification test. After three years, consulting teachers and evaluating teachers returned to teaching full-time to continuously work on their practice. Donaldson realized the importance of multiple evaluators with content knowledge and expertise to provide diverse feedback to teachers (Donaldson, 2010). The changes made to the Cincinnati Teacher Evaluation System, including ways of providing feedback, incorporating the Danielson framework, and including various experts, had a positive impact on overall instruction and student achievement (Kane, Taylor, Tyler, & Wooten, 2010). In conclusion, developing a framework that includes common vocabulary and that utilizes experts to provide teachers with feedback helps increase teacher effectiveness and student achievement.

**Teacher collaboration and goals.** In order for teachers to grow through the evaluation process, they must have ownership in the process by engaging in collaboration and goal setting activities (Danielson, 1996; Darling-Hammond, 2014; Stiggins & Bridgeford, 1985; Stronge & Helm, 1992). Wise, Darling-Hammond, McLaughlin, and Bernstein (1985) looked across the country for effective teacher evaluation systems. Effective evaluation systems increase communication between the administration and teachers, which improves the overall quality of work (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1985). The few evaluation systems that they felt were effective included the following criteria: incorporating teachers’ goals, providing meaningful feedback, and collaborating with other teachers. An effective evaluation system focuses on collaborative opportunities (Darling-Hammond, 2014). Because teachers and evaluators
found the teacher evaluation process to be a superficial and meaningless exercise, the modern school reform movement in 1983 focused on shaping a more meaningful and collaborative evaluation process to improve teacher quality. In order to improve the teacher evaluation process, teacher quality became the focus of policy change. Supporters believed that the teacher evaluation process needed to include collaborative learning environments for teachers to engage in professional dialogue. The modern school reform movement attempted to link teacher evaluation with coaching and collaboration (Danielson, 2001).

Race to the Top (RTTT) federal funding opportunities and No Child Left Behind (NCLB) flexibility waivers required school districts to implement new teacher evaluation systems. Requirements within the new systems included four-point rating systems, numerous observations, and the practice of using student test scores to rate teacher performance. The compilation of data impacted tenure, contracts, advanced certification, and dismissal of teachers. These governmental financial incentives led many states to reorganize their evaluation systems. Solely changing the systems to improve teaching was not the solution because, “we need to create and sustain productive, collegial working conditions that allow teachers to work collectively in an environment that supports learning for them and their students” (Darling-Hammond, 2014, p. 2).

Noticing the continuous challenges faced by the teacher evaluation system, the National Association of State Boards of Education recognized the need for better recruitment of teachers, preparation programs, and long-term professional development (Darling-Hammond, 2014). All areas needed to be aligned in order to improve the overall education system (“Moving beyond compliance,” 2015). The system included common standards across states, performance-based assessments, and professional development linked to the
evaluation (2014). The state licensure systems needed to align with the evaluation. Interstate Teacher Assessment and Support Consortium (InTASC) Standards were adopted in more than 40 states, aligned with a few pilot programs. These programs included Performance Assessment for California Teachers (PACT), the Oregon Teacher Work Sampling (TWS) and the Teacher Performance Assessment (edTPA). Darling-Hammond stated, “Because yearly district evaluations are based on the same standards as licensing assessments, teachers can continue to work on their practice coherently throughout their careers” (2014, p 5). In order to develop an authentic evaluation, teachers need to work collaboratively and focus on aligning their goals with the evaluation criteria (Danielson, 1996; Stronge & Helm, 1992).

The center of education is teaching and learning, which cannot be accomplished without high quality teachers and goal-focused assessments (Stronge & Helm, 1992). “If goal establishment and goal accomplishment are fundamental to success, the evaluation system should reflect this orientation” (p. 175-180).

An effective teacher evaluation process entails collecting data about performance, planning for improvement and a fair system of accountability. It also needs formative and summative assessment components. The summative portion is accountability focused, while the formative side concentrates on improvement. The goals, in all areas of the assessments, must be mutually beneficial for the teacher and the school district. Formative assessment identifies the teachers’ specific strengths and weaknesses, develops professional goals based on that information, and promotes self-regulation during the assessment process (Stiggins & Bridgeford, 1985; Nicol & Macfarlare-Dick, 2006)). “Formative evaluation systems, inherently sensitive to teachers' needs and goals, can be a vital step in strengthening instructional effectiveness nationwide” (p. 96). Effective evaluative tools reflect teachers’
goals and provide multiple collaborative opportunities for teachers to receive helpful feedback (“Moving beyond compliance,” 2015).

**Standard-based assessments.** Alternative evaluation systems have also been implemented in various school districts in an effort to study their impact. Milanowski (2004) studied an alternative evaluation focused on a standard-based assessment, based on Danielson’s Framework for Teaching (Danielson, 1996). This system contained detailed standards that reflected effective teaching strategies, used multiple levels of performance, and contained a detailed rating scale. Within this system, the evaluators were trained and collected extensive amounts of data about the teachers, which included multiple observations and artifacts, in order to accurately assess teacher performance and the effects on student learning. In order to determine the impact of the standard-based system, Milanowski (2004) implemented the system within the Cincinnati Public Schools (CPS), an urban district with a high-level of low-income students who participate in free or reduced-lunch. The teachers in the CPS received scores in four areas: planning and preparation, classroom management, teaching for learning, and professionalism. After implementing the system for two years, Milanowski concluded that the scores provided to the teachers had moderate validity. The scores showed which teachers’ students had “higher than expected levels of achievement to a degree higher than chance” (p. 49). As a result, Milanowski determined that scores from a detailed teacher evaluation system can be directly related to student achievement and used to decide pay differentiation and performance remediation within school districts.

Chung and Pecheone (2006) also explored a standard-based assessment that could impact evaluation systems. They discussed the development of the Performance Assessment for California Teachers (PACT), a standards-based performance assessment used to make
credentialing decisions in teacher preparation programs. Chung and Pecheone used a two-year pilot to examine the reliability of the assessment and the teacher events (TEs), which included teacher plans, teacher artifacts, student samples, video clips, and other components. They studied how accurately the PACT and TEs assessed teacher performance and the teachers’ impact on student learning. The research illustrated that the PACT and TEs act as valid measures for the purposes of teacher credentialing and improving credentialing programs. One concern for maintaining such an assessment is the long-term cost, yet without an assessment focused on teacher performance, Chung and Pecheone expressed concerns that teachers would be solely judged by student scores on standardized, multiple-choice tests. “In this current policy environment, it appears that it will not be long before states begin to develop alternative teacher credentialing routes that depend solely on test scores and bypass any clinical demonstration of teaching skill” (p. 32). Chung and Pecheone believed that in order to measure teacher performance as well as student learning during the evaluation process, a teacher assessment system with multiple measures needed to be implemented.

Competency-based assessment is a type of assessment that sets a certain level of competency in order to determine someone’s professional ability (Gonczi, Hager, & Athanasou, 1993). If a specific level is determined for all new employees, then they must all meet the standards of performance. Standards can change depending on the position. For example, there are different standards for entry-level positions versus roles with more experience. “Under a competency-based assessment system, assessors make judgements based on evidence, about whether an individual meets criteria specified in the profession's competency standards” (p. 1). A professional is considered competent when they demonstrate the necessary qualities to meet the set standards. “Competence is focused on
performance of a role or set of tasks” (p. 5). Adequate evidence must be collected prior to making any judgement about someone’s professional ability. Examples of evidence include tests, assignments, logs, and observations. When using competence-based assessment, the individual’s performance is compared to pre-developed criteria, not compared to groups of people as seen in norm-referenced assessments.

Teachers need to look at assessments through a different lens and see them as helpful and insightful (Chudowsky & Pellegrino, 2003). All assessments should ultimately provide information that positively impacts student learning and helps students succeed in school (Chudowsky & Pellegrino, 2003; Shepard, 2000).

**Multiple measures.** Multiple measures are needed to accurately assess student learning and teacher performance (Haertel, 1986; Steele, Hamilton, & Stecher, 2011). Student performance assessments cannot be the only indicators used to determine either student or teacher growth. Many times student performance assessments are encouraged for summative teacher evaluations, in order to control teacher quality and measure student performance (Chung & Pecheone, 2006). Student achievement tests reveal the teachers’ successes in increasing overall cognitive learning in every child.

It has been argued that student achievement depends on multiple factors, many of which are out of the teacher's control, and also that published, standardized tests are unlikely to match the learning objectives of a particular teacher at a particular time (Millman, 1981, p.154).

Standardized tests do not include all learning objectives, which lead to a false indication of student overall learning. The lack of reliability of standardized tests to reflect student
learning leads the tests to be poor indicators of overall teacher performance (Millman, 1981; Steele, Hamilton, & Stecher, 2011).

Multiple pieces of data must be implemented to accurately assess teacher performance (Danielson, 1996; Haertel, 1986; Gonczi, Hager, & Athanasou, 1993; Steele, Hamilton, & Stecher, 2011; Stiggins & Bridgeford, 1985). Formative evaluations need to include feedback from students, other teachers, administrators, and the person receiving the evaluation and should also include multiple observations. The Consortium for Policy Research in Education studied whether teacher collaboration, focused on student data and data compiled from observations, impacted instruction and student achievement (Darling-Hammond, 2014). The study assigned teachers to two groups. The control group only discussed student data, while the other group talked about student data in conjunction with data compiled from observations. The group that discussed both student and observation data led to students having significantly larger learning gains and teachers who integrated more instructional changes. Focusing solely on student test data may seem easier for school districts, yet “unskilled use of this kind of data can have damaging ramifications due to the misevaluation and potential loss of good teachers and the incentives for teachers to avoid the neediest students” (p. 5). In order to measure teacher performance, multiple measures are needed.

A number of factors impact student learning, including various background characteristics, such as student mental ability, student motivation, and parental support. School variables, including class size, teach load, teacher training, material access, school facilities, and overall school climate also affect student learning (Haertel, 1986; Drago-Severson, 2012). Student tests that are linked to evaluation processes tend to reflect students’
socioeconomic status (SES) more than teacher instruction. These tests appear to measure the SES found within a school rather than the actual efforts and effectiveness of teachers. Testing does not accurately reflect teacher performance (Popham, 2007; Steele, Hamilton, & Stecher, 2011). “Standardized tests tend to be unsuitable for measuring educational achievement as distinct from student competence, because they sample broad subject domains and are unlikely to match closely with the curriculum in particular classrooms at particular times” (Haertel, 1986, p. 50). In order for standardized tests to accurately assess teacher performance and be used as part of the evaluation process, test content must parallel the set curriculum and variables need to be eliminated.

In summary, multiple measures must be considered to accurately assess teacher performance (Danielson, 1996; Haertel, 1986; Gonczi, Hager, & Athanasou, 1993). “Evaluations should be based not only on value-added estimates that use student test scores but also on other data sources, such as observations of teachers' classroom practices and evidence of their contributions to their schools” (Steele, Hamilton, & Stecher, 2011).

Feedback. Evaluators need to provide teachers with accurate feedback in a timely manner. Teachers and evaluators must share and agree upon common language in regards to effective teaching. Evaluators need to visit classrooms frequently. Unannounced mini-observations allow the principal to stay focused on observing during busy days and provide accurate data on teacher performance. Collecting data from mini-observations allows the evaluators to have an array of snapshots that lead to a better understanding of the teacher and the students (Marshall, 2005). After every observation, teachers must receive punctual, high-quality feedback. Principals must frequently visit classrooms and provide prompt feedback (Mendels, 2012). Marshall (2005) encouraged evaluators to use data from several mini-
observations, instead of less-frequent formal observations, to write-up an overall evaluation. Teachers need to also provide evidence of student learning as part of the evaluation process.

When a professional learning culture develops in a school, teachers begin to provide accurate feedback to peers and become less defensive. “In the context of a school’s collaborative culture, walkthroughs provide an effective structure for this kind of dialogue” (Ginsberg & Murphy, 2002, p. 5). Also, by engaging in short, frequent, and focused walkthroughs, evaluators observe teacher patterns in curriculum and instruction (Downey, Steffy, English, Frase, & Poston Jr., 2004).

New York City Public Schools decided to increase unannounced observations, using the Danielson rubric and the language associated with the rubric, and implementing a four-point scale (Marshall, 2014a). The district expected evaluators to score teachers after every observation. However, as mentioned in “How the Danielson Rubric Could be More Effective for City Teachers and Principals,” scoring during each observation can take away from the authenticity of the evaluation and the ability of an evaluator to organically watch the instruction of the teachers (Marshall, 2014a). The best way to observe a classroom is to keep your “head up, listen carefully to teacher/student interactions, scan what’s on the walls, and look over students’ shoulders to assess the instructional task” (p. 1). When evaluators solely focused on scoring the rubric, overall feedback declined because they did not have the opportunity to watch the classroom and understand the dynamics. Evaluators should write down brief notes and discuss the information during a post-visit conference. This strategy allows the evaluator to provide feedback that is collaborative and less about judgement. The strongest coaching moments happen during post-observations, which New York City did not require (2014). The feedback paperwork associated with each observation reduces the
amount of time an evaluator spends in classrooms. Unless a teacher is truly struggling, providing rubric scores after every observation can be excessive.

In order for teachers to value the provided feedback, evaluators need to be present in the classroom at least every month and focus on one or two objectives at a time. “Teachers getting feedback on 10-15 areas of a rubric can be overwhelming, especially if the feedback is critical” (Marshall, 2014a, p. 2). Also, Danielson’s rubric was not originally designed for feedback during mini-observations or evaluations, and research does not support the use of the rubric, or similar measurements, in this manner. “Because teaching is complex, it is helpful to have a road map through the territory, structured around a shared understanding of teaching” (p. 2). The Danielson framework focused on providing equitable access to all students to achieve success, developing common terminology among teachers to hold each instructor to a high standard, and guiding collaborative opportunities (Danielson, 1996). The Danielson rubric has since been used to gauge the needs of the school’s teachers, administrators, and students, and focused on the relationships between variables both inside and outside of the classroom (Danielson, 2002).

In order for teachers to grow, feedback must accurately depict the teachers’ practices and focus solely on transmission skills (Darling-Hammond & McLaughlin, 1995). According to Marshall (2014a), evaluators also need time to reflect on their practices and develop new ways to effectively establish a culture of learning. “Being reflective can help reinforce cultural patterns that are positive and transform those that are negative or toxic” (Deal & Peterson, 2011, p. 10). The school culture also impacts how teachers and administrators identify with the school, as well as affect their overall productivity and commitment towards achieving the core values and goals (Deal & Peterson, 2011). High
quality evaluators must visit classrooms often, understand instruction, know how to evaluate instruction, and effectively share feedback in an authentic way. Marshall (2014a) supported the idea that evaluators should operate in a fishbowl-like environment, where multiple administrators share their ideas about an observation and collaborate on the findings. Then, the group asks the teacher their ideas about the feedback. By engaging evaluators in a collaborative exercise that provides feedback, they are reminded about how it feels to be evaluated (Marshall, 2014b).

Evaluators tend to provide vague, inflated feedback (Donaldson, 2010). Jones, Bergin, and Murphy (2018) recently completed a study that focused on collecting information from 340 teacher evaluators. The research focused on the evaluators’ motivations for giving certain feedback to specific subgroups of teachers. The data showed that evaluators often provide inflated ratings to veteran teachers who were close to retiring as well as new teachers. Evaluators indicated that they did not see good reason for providing challenging data to teachers who were leaving the field and did not want to discourage teachers just entering the field either. These types of tendencies led to inaccurate data being shared with certain subgroups of teachers by evaluators.

The Consortium on Chicago School Research at University of Chicago (Sartain, Stoelinga, & Brown, 2011) studied the current challenges with the teacher evaluation system and the reasons behind so many reforms across the country. The two main factors for redevelopment within the system are teachers not currently receiving quality feedback and the evaluation not accurately differentiating between high- and low-quality teachers. Ninety-three percent of the teachers in Chicago were rated in the two strongest categories of effectiveness, while just 0.3% of teachers were determined to be in the lowest performing
category. These types of inflated rankings did not give teachers the feedback they needed to grow and learn, and also failed to accurately identify low-performing teachers.

In order to improve the system, policymakers developed the Race to the Top agenda, which made districts change their teacher evaluation systems in order to receive funding. The grant money is intended for schools that develop innovative plans for improving student performance and increasing overall school productivity (U.S. Department of Education, 2009). According to Sartain et al. (2011), the U.S. Department of Education provided $4.35 billion in Race to the Top funds. In 2010, in response to the Race to the Top requirements, Illinois developed the Performance Evaluation Reform Act, which asked all schools to use a standard-based evaluation coupled with student achievement data and a rubric for all observations by the 2015-2016 school year to increase teacher feedback. Currently, this system is still in place even though Illinois did not receive the Race to the Top funding. Race to the Top funding requirements and findings from the Excellence in Teaching Pilot in 2008 impacted the format for the current system, which included a standards-based observation rubric (Sartain, Stoelinga, & Brown, 2011). Currently, recipients of the grant money have developed achievable plans towards education reform, which ultimately transforms schools (U.S. Department of Education, 2016).

Before this pilot, Illinois used a traditional system for 30 years with a checklist, one observation, one evaluator, and no connection to student data. The pilot concentrated on enhancing teaching and learning professional development for teachers and additional training for principals. Forty-four elementary schools participated in the pilot. Evaluators used the Danielson Framework during observations which occurred twice during the school year. Each observation involved a pre-conference, classroom observation, alignment with
rubric by the evaluator, and a post-conference. Half the elementary schools began the program in 2008 and the remainder in 2009. Schools were picked randomly and placed in one of the two groups. Generalizability increased due to that random selection. The Consortium on Chicago School Research (CCRS) randomly selected teachers within the elementary schools. Reliability increased because two evaluators were used to rate each teacher. Validity came from implementing observations of teachers who had value-added measures from Chicago Public Schools (Sartain et al., 2011). CCRS compared the value-added measures and the rating of teachers for all four domains of the Danielson Framework. Through qualitative research, CCRS discovered the principal and teacher perceptions of the tools and the challenges and successes of the system. They interviewed teachers and principals and engaged in case studies in eight schools, selecting schools with various levels of principal engagement. The findings indicated that principals ranked teachers consistently at the lower or middle of the scale, in contrast to the MET Report that indicated that observations are unreliable (Sartain, Stoelinga, & Brown, 2011; Resmovits, 2013). Most of the time, the two evaluators agreed on the scores. In the evaluations, teachers frequently received lower ratings regarding instruction rather than classroom environment. Data from interviews indicated that principals wanted more training and support within instructional leadership, and that teachers had limited understanding of the evaluation process. Findings also indicated that only 10% of questions asked of teachers were high-level, and principals seemed to do most of the talking during conferences. In summary, the pilot project illustrated the need for more reflection on practice and training to better understand all of the components of the evaluation system and to provide more meaningful feedback (Sartain et al., 2011).
In conclusion, evaluations and assessments of teachers are multi-faceted. They must include components of accountability and reflection, collaboration and goal setting, and feedback. In order to accurately evaluate teachers, multiple measures must be included in the evaluation process (Haertel, 1986; Gonczi, Hager, & Athanasou, 1993).

**High-stakes testing and pay-based incentives.** State tests are frequently used to assess teacher performance and student learning. Kimball, White, Milanowski, and Borman (2004) studied the relationship between a standards-based teacher evaluation system and student results on state tests. Using a value-added framework, they found an initially positive relationship between the two components, yet the coefficients were not statistically significant in all situations. Potential reasons contributing to the lack of consistency in the data included broadly written standards which made it difficult to assess teacher performance and the context of the teacher evaluation. First, the same standards and evaluation instrument were used for teachers teaching various grade levels and subjects. The evaluation system focused very little on actual instructional content. “This evaluation system is generic with respect to instructional content” (p. 71). Second, the evaluation system was low stakes. The system focused more on increasing overall morale and little on differentiating teacher performance. Regarding the direct impact a standards-based teacher evaluation system has on student state assessments, the study is inconclusive.

Haertel (1986) indicated that school variables including class size, teacher training, material access, and school facilities, as well as student factors such as student backgrounds, student motivation, and parent support, all have an impact on student learning. Student tests that are linked to evaluation processes tend to reflect students’ socioeconomic status (SES) more than teacher instruction. These tests measure the SES found within a school rather than
the efforts and effectiveness of teachers. Testing does not accurately reflect teacher performance (Popham, 2007). In order for standardized tests to accurately assess teacher performance and be used as part of the evaluation process, test content must align with the set curriculum and the variables that could skew the data must be eliminated. “Standardized tests tend to be unsuitable for measuring educational achievement as distinct from student competence, because they sample broad subject domains and are unlikely to match closely the curriculum in particular classrooms at particular times” (Haertel, 1986, p. 50). Because standardized tests focus on large amounts of curriculum, the tests become sensitive to individual student differences more than classroom instruction and exposure to material. Teacher effectiveness cannot be judged by standardized tests unless multiple variables are taken into account (Haertel, 1986). Teacher perspectives are important when deciding how to use test data as part of the evaluation process, yet this research does not reflect teacher perspectives.

The design and improvements of assessments should focus on enhancing student learning, but unfortunately recent external assessments have been used in ways that demotivate students and provide little insight into student learning. (Shepard, 2000) “High-stakes accountability teaches students that effort in school should be in response to externally administered rewards and punishment rather than the excitement of ideas” (p. 8). High-stakes testing can actually decrease student learning, yet show an increase in testing performance (Shepard, 2000). Chudowsky and Pellegrino (2003) further asserted that assessments need to change and align with learning objectives to enrich student learning and to guide teaching. Assessments must encourage authentic learning connected to real-world experiences, embody critical thinking and problem-solving skills, and inspire students to
embrace difficult learning challenges. To achieve this degree of transformation, the assessments need not only to become more informative and better linked to learning objectives, but the meaning of those evaluations also needs to change. Teachers and students need to look at assessment through a different lens and see assessment as helpful and insightful. All assessments should provide information that positively impacts student learning and helps students succeed in school (Chudowsky & Pellegrino, 2003; Shepard, 2000).

The Measuring Effective Teachers’ Report (MET), published by the Bill & Melinda Gates Foundation (2013), found that test scores are more reliable than observations when measuring teacher effectiveness. Indeed, teacher effectiveness can be measured. They came to this conclusion by conducting a randomized experiment using observation, testing, and student surveys equally. Teachers whose students improved the first year also made gains in the second year on standardized tests. This finding caused a fundamental shift in an education community that provides increased salary and benefits to teachers with more years of experience and indicated that teachers may be hired and fired based on performance rather than seniority. The Race to the Top encouraged schools to implement testing as part of the evaluation system (U.S. Department of Education, 2009; U.S. Department of Education, 2016). The MET Report also indicated that observation is extremely unreliable, though using additional evaluators increases the reliability slightly. "The MET findings reinforce the importance of evaluating teachers based on a balance of multiple measures of teaching effectiveness, in contrast to the limitations of focusing on student test scores, value-added scores or any other single measure," stated Weingarten from the American Federation of Teachers Union (Resmovits, 2013, p. 3).
Many pay-based incentive programs emphasize student standardized tests, yet the pay-based incentive program should focus on increasing teacher knowledge and skills. The professional benchmarks vary depending on the teacher’s experience. Districts encourage beginning teachers to take PRAXIS II and the Danielson Basic to test content knowledge and pedagogical understanding. The Danielson assessments focus on planning and preparation, classroom environment, instruction, and professional responsibilities. Intermediate teachers focus on the PRAXIS III, which centers on overall practice, and the Danielson Proficient assessment. Districts encourage advanced teachers to get their masters degrees, to become nationally certified, and/or complete Danielson Advanced. By assessing teachers based on knowledge and skill, and by providing a pay-for-performance incentive, teacher performance and accountability increases. This program promotes standards-based education reform and enhances student learning. The pay-based incentive programs must not solely focus on student standardized tests (Odden, 2000).

The pay-based incentive program currently under development in Douglas County, Colorado, focuses more on teacher skills and performance, and less on student standardized tests. Douglas County Schools changed their compensation and evaluation system to encourage recruitment of the best teachers. The changes included developing a differentiated pay scale that encompassed ranking teachers as highly effective, effective, partially effective and ineffective, which provided teachers with more money based on their ranking, thus rewarding effective teachers. The school system also assigned each position a certain pay band, for example math teachers. If effective math teachers are difficult to find within the market, Douglas County Schools encourages more applications by increasing the salary. This experiment with Douglas County Schools will hopefully continue to attract highly
compete educators, motivate teachers already in the school system, and increase student learning ("The school system’s ideas," 2013).

**Value-added model.** The teacher evaluation system has been analyzed by looking at teacher contributions to student learning, as well as the effects of the value-added models (VAMs), which focus on student test-scores across multiple years, while integrating factors that impact student achievement. VAMs focus on the idea that student achievement gains are a reflection of teacher effectiveness. “Student learning is measured well by a given test, is influenced by the teacher alone, and is independent from the growth of classmates and other aspects of the classroom context,” which are not found in research (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012, p. 8). Many factors beyond the teacher impact student achievement including class sizes, student needs, peer culture, and home environment. The problems with VAMs include inconsistency, a failure to consider other influences that can impact student achievement, and the ways the students assigned to teachers affect the evaluations. “VAMs wrongly attribute to teachers other influences on student performance that are present when the teachers have no contact with students” (p. 9). VAMs can be used in a way that misjudges teacher performance and leads to untrue generalizations, which ultimately skews educational decisions (McCaffery, Lockwood, Koretz, & Hamilton, 2003). This statistical technique compiles data over numerous years to measure effects of teachers and schools, which districts use as accountability measures.

Effective evaluation systems, in contrast, include numerous observations, trained evaluators with expertise, multiple pieces of data, and prompt and meaningful feedback. Standards-based evaluations, such as the National Board Certification and new teacher performance assessments, encourage teachers to measure their practice and results, utilizing
videotaping, providing artifacts, and receiving student feedback (Darling-Hammond et al., 2012).

**Conclusion**

After reading years of research on teacher evaluation processes, there appears to be a gap in the research when focusing on teachers’ perspectives, especially on the observation portion of the teacher evaluation system. Evaluators have limited time to observe teachers during the school year due to demanding schedules (Ginsberg & Murphy, 2002). In order for evaluators to accurately assess teachers, they need to spend adequate time observing teachers’ practices (Resmovits, 2013). In order for the classroom observations to effectively impact student learning, teachers need to value the tool and use the feedback to create future goals (Danielson, 1996; Danielson, & McGreal, 2000).

This research project focused on asking teachers their perspectives about the successes and challenges of the current observation system. By listening to the perspectives of teachers, the evaluation process can become a more engaging process, leading to a more active evaluation document (Heneman & Milanowski, 2001). This research project set out to answer such questions and to learn about teachers’ perspectives on the observation system.
Chapter 3: Methodology

Some teacher perspectives have been incorporated into the development of teacher evaluation tools. However, this study focused specifically on discovering teachers’ perspectives on the classroom observation portion of evaluations, a few years after the teacher evaluation system was implemented. The purpose of this study was to understand how teachers perceive the classroom observation of the evaluation system, and to discover the challenges and successes of the evaluation system that were in place at the time of this study from the teacher perspective.

Methodological Framework

Phenomenology. In order to understand the perspectives of the teachers in regard to the observation process within the evaluation system, I used the approach of phenomenology to conduct this qualitative study.

Qualitative research focuses on finding multiple realities and truths to explain a research topic (Creswell, 2013). Qualitative research is appropriate when the research questions seek to understand how, what, or why things occurred (Mertens, 1998). The design of this qualitative study was flexible and constantly evolving.

Phenomenology focuses on understanding a common lived experience, or phenomenon (Creswell, 2013). A lived experience is “experience as we live through it and recognize it as a particular type of experience” (van Manen, 1990, p. 177). The idea of phenomenon focuses on exploring concepts. A phenomenological study concentrates on a number of individuals and commonalities found within their experiences of a phenomenon. For example, people around the world experience anger. This universal feeling can be studied using phenomenology, because it is a common lived experience.
Specifically in education, the concepts of teacher evaluations can also be studied using phenomenology because this group of individuals has all experienced the phenomenon. The main objective of a phenomenological study is to look for patterns among the individuals and to compile the experiences into a unique description of the nature of the lived concept, or the “whatness of things” (van Manen, 1990, p. 177). Researchers identify a common meaning from lived experience, collect data from individuals who experienced the phenomenon, and create a multifaceted description for all of the participants (Creswell, 2013). The description encapsulates that whatness of the experience. It also focuses on specifically how the participants experienced the phenomenon.

Many researchers, including Hüsserl, the founder of phenomenology, debated the structure and use of phenomenology for years. Common characteristics within the philosophy include the idea of a common lived experience, individuals consciously engaging in the experience, and a need to craft descriptions of the experiences by each individual (Moustakas, 1994). Because phenomenology focuses on conscious lived experiences, this methodology is commonly used in education (van Manen, 1990).

**Epoché.** Specifically, phenomenological studies focus on a few basic philosophical ideas (Stewart & Mickunas, 1990). The first philosophical idea is “epoché,” which Hüsserl discovered. Epoché, or bracketing, is “suspending one’s various beliefs of the reality of the real world in order to study the essential structures of the world” (van Manen, 1990, p. 175). A researcher separates their own ideas in regard to the phenomenon being studied in order to create a blank slate. The idea of bracketing is to create, as much as possible, an environment with limited judgement by the researcher toward the participants and their experiences. The researcher tries to bracket her experiences by analyzing and discussing her own experiences.
with the phenomenon (Creswell, 2013). Giorgi (2009) stated that bracketing does not allow researchers to forget their personal experiences, but does help separate their own experiences to better understand other individuals’ experiences.

**Intentionality of consciousness and noema.** The second philosophical idea highlights the *intentionality of consciousness*. *Intentionality of consciousness* refers to the intended relationship between the *act of consciousness* and the *object of consciousness* (Moustakas, 1994). The way individuals perceive objects depends on their own consciousness; thus reality consists of individuals developing ideas about objects through their consciousness. In summary, individuals have both subjective and objective experiences, leading many to place phenomenology between the worlds of qualitative and quantitative research (Creswell, 2013).

According to Moustakas (1994), Hüsserl discussed the idea of consciousness even further, coupling it with intentionality. For individuals to create meaning, they must perceive, remember, judge, etc. Through the process of creating meaning, the consciousness of individuals intentionally steers them towards specific thoughts and ideas.

Hüsserl also examined the idea of *noema*, which directs consciousness towards specific ideas. Through *noema*, individuals develop meanings within experiences and develop clarity of what is present in their consciousness. Hüsserl also discussed the idea of *noesis* which “refers to the act of perceiving, feeling, thinking, remembering or judging - all of which are embedded with meanings that are concealed and hidden from consciousness” (p. 69). *Noesis* expresses intentionality. There are differences between the way an individual’s mind perceives an object and the actual object. The perception of an object is the *noema*, while the meaning that comes from the perception of object is the *noesis* (Moustakas, 1994).
Through the relationship between the *noema* and the *noesis* develops the *intentionality of consciousness*. When involved in a phenomenological study, researchers must consider the meanings developed through the consciousness of the individual, separate from the means they themselves have developed.

**Variations of phenomenology.** Phenomenology includes two variations: *hermeneutic phenomenology* and *transcendental phenomenology*. *Hermeneutical phenomenology* focuses on the interpretation of life within the lived experiences. Van Manen (1990) discussed this type of phenomenology as six separate stages including discovering the phenomenon, reflecting on themes which describe the essence of the lived experience, and writing a detailed description about the phenomenon. “Phenomenology is not only a description, but it is also an interpretive process in which the researcher makes an interpretation of meaning of the lived experiences” (Creswell, 2013, p. 80).

*Transcendental phenomenology* emphasizes more on description than interpretation, in order to achieve a natural view, which is difficult to accomplish. Moustakas (1994) discussed how the steps of *transcendental phenomenology* include identifying the phenomenon, bracketing, and collecting data. Researchers collect data from individuals who experienced the phenomenon. Then, researchers analyze the data, develop themes, and compose a description, which explains what and how individuals experience the phenomenon. Researchers also develop both textural and structural descriptions to grasp the essence of the participants’ experiences. The textural experience looks at the actual experiences of the participant, while the structural descriptions focus on the conditions and setting of the experience. Together, these descriptions allow the essence of the experience to
emerge (Creswell, 2013). Both variations of phenomenology focus on the lived experience of a group of individuals.

During this study on teachers’ perspectives of the evaluation system, I used phenomenology, a strong research approach for understanding the shared experiences of many teachers. As Merriam stated, “From the philosophy of phenomenology comes a focus on the experience itself and how experiencing something is transformed into the consciousness” (p. 24). I wanted to learn about teachers’ experiences with evaluation systems and how their experiences impact their teaching and life. I sought to understand how the day-to-day experiences involving the evaluation process made an impact on the teachers’ conscious experiences. By engaging in bracketing, I put aside my own lived experiences so my interpretations of the teachers’ lived experiences were consistent and not broken. In order to bracket my experiences, I first analyzed my own biases, prejudices, and assumptions. To portray an authentic description of teachers’ perspectives and to control the outcome judgement, my own experiences needed be isolated.

By understanding the teachers’ common experiences, I worked to adjust and improve future practices and policies. More specifically, I used transcendental phenomenology, an approach that concentrates on descriptions and insights from participants in addition to how the meanings of objects form through consciousness. I came away from this study with a better understanding of what it feels like to experience the evaluation process from the teacher viewpoint. I sought to understand how the teachers interpreted and valued the teacher evaluation process. Phenomenology allowed me to understand the essence of teachers’ experiences within the evaluation system.


Research Activities

Epoché and transcendental-phenomenological reduction. Because of my direct experience with the phenomenon, I participated in epoché, or bracketing. Epoché is when a researcher sets aside her biases and assumptions to take an open perspective towards the phenomenon. “The world is placed out of action, while remaining bracketed” (Moustakas, 1994, p. 85). Engaging in epoché is an experience within itself. It challenges the researcher to see all information as equal and as new knowledge. Researchers attempt to place experiences into their consciousness, as though to see everything as new (Moustakas, 1994). Moustakas (1994) explained that, to engage in epoché, researchers must go to a quiet place and think about an experience then attempt to think about the experience again and again until they feel no biases towards the experience. Clearing the mind takes practice, but when the mind is cleared, researchers gain new knowledge as well as a new beginning for understanding concepts. “The epoché is a way of looking and being, an unfettered stance” (Moustakas, 1994, p. 85). Moustakas encourages reflective meditation for clearing the mind because a researcher may not be able to achieve bracketing when a specific experience is extremely deep or emotional. In order to ensure quality data collection and analysis, I attempted to bracket my experiences with the phenomenon.

As bracketing is a difficult task for researchers, I achieved this state by first analyzing my position as a researcher and my current assumptions and biases. As a researcher, I have been both an administrator and teacher. Serving in both of these roles throughout my career provided me with a different lens when analyzing the teacher evaluation process. As a teacher, I knew what it felt like to have an evaluator come into my classroom and provide feedback. Sometime, due to the lack of presence in my classroom, I felt the feedback was
useless. When an evaluator dedicated time to understanding my classroom environment and the students, I found the feedback helpful. As an administrator, I understood the challenge of finding time to evaluate teachers. As visiting classrooms often was always my goal, other fires tended to flare up during the school day that immediately took my attention. Even though I planned my days to visit classrooms, often my days turned into something completely unplanned. Before engaging in this research, I assumed other teachers felt similar to me, which included feeling that teachers’ perspectives should guide the development and implementation of the evaluation system, observations should occur frequently, and evaluators should be content experts in order to provide useful feedback. My biases included thinking that the process should focus more on developing goals that incorporate the needs of both the teacher and the school district. I thought both parties, the evaluator and teacher, needed to actively engage in the process and not expect the other party to fulfill every component. For example, I heard some teachers at my last school state that the evaluation process is not engaging, however, they too need to try to make the process useful. I engaged in reflective meditation for multiple days. Bracketing occurred before initial interviews, in order to approach the interviews with an unbiased presence.

I engaged in epoché to set aside my biases and assumptions, so I could clearly focus on consciousness itself. “Epoché requires a new way of looking at things, a way that requires that we learn to see what stands before our eye, what we can distinguish and describe” (Moustakas, 1994, p. 33). Following epoché, I engaged in transcendental-phenomenological reduction, which helped me “to grasp the structural essences of experience” (Moustakas, 1994, p. 35). Transcendental-phenomenological reduction takes the researcher back to the essence of the experience to revisit an experience in order to better understand its original
structure, which allows the researcher to separate the phenomenon. Van Manen (1990) defined essence as “the true being of a thing” (p. 177). This separation allows the researcher to better understand the essence of the phenomenon. “When we perceive straightforwardly, we focus on the object itself and not the perceiving experience” (Moustakas, 1994, p. 91). Transcendental-phenomenological reduction works to help a researcher concentrate solely on the essence of the phenomenon and listen to the participants through their own meanings and perceptions of that phenomenon. “Ultimately, through transcendental-phenomenological reduction, we derive a textural description of the meanings and essence of the phenomenon, the constituents that comprise the experience in consciousness from the vantage point of an open self” (p. 34). By engaging in epoché and transcendental-phenomenological reduction, I absorbed new knowledge from participants through a pure consciousness that allowed me to better understand the essence of the phenomenon being researched.

Understanding positionality. Before research began, I reflected on my own biases and the threats that could result from those biases and possibly affect the conclusions that arise from my research, impacting the validity. This process ultimately led me to analyze each step of the research process and create strong inferences about teachers’ perspectives of the teacher evaluation process. I wanted readers of the final research to remain confident that I analyzed the data from many viewpoints and considered my own personal biases and research threats when analyzing the data. By analyzing my own biases and research threats, I minimized the challenges that could occur during the research process (Corbin & Strauss, 2008).

Limitations. I began by analyzing my limitations. I had biases resulting from having been both an insider and an outsider during the teacher evaluation process. I associated as an
insider because I was evaluated as a teacher by professors at University of Notre Dame, as well as principals, assistant principals, and peers in private and public schools. I understood the process first-hand. I also identified potential biases from being an outsider, having served as a school administrator and evaluator. I knew the challenges of the teacher evaluation process through the lens of both a teacher and administrator. Seeing this process from both the perspective of a teacher and evaluator led me to have strong opinions about what the evaluation process should entail. Because of my insider views and perspectives, sometimes I assumed I knew how other teachers felt about the process. I needed to constantly remind myself to stay open-minded, while learning new perspectives. When selecting a sample of teachers, I gravitated towards people I knew and, most likely, had shared my opinions in the past. However, I realized that I needed to continuously assess my sample of participants to ensure they closely represented the population of teachers in the city studied.

During the completion of the research, I was somewhat of an outsider to the evaluation process. I was not working in the public education system at the time, and therefore did not know how it felt for state testing data to be used as a primary measure of effective teaching by district offices and school administrators conducting the evaluations. My lack of involvement also led to potential biases. I assumed that I understood what the evaluation process entailed, but I did not completely comprehend the challenges involved in the teacher evaluation process that had evolved due to policy changes in the years immediately previous to the research.

A possible threat to my research was the timeline and financial constraints. My involvement in the doctoral program impacted my family financially, imposing a narrow timeline on my research. Prolonged engagement in the field of study allows researchers to
learn the culture, develop trusting relationships, and check for misinformation. Specifically, engaging in phenomenology takes additional time, because this approach involves multiple interviews. Finding time to meet multiple participants and engage in interviewing takes time. I did not want these important aspects of the research to be negatively impacted by my timeline or financial challenges. Through constant reflection, I worked on making sure that I did not rush to a conclusion, but allowed adequate time for saturation of the data to occur. Saturation occurred when the information from the data fully developed and became repetitive (Creswell, 2013). Also, I did not want the validity of my research to decline because of my financial and timeline constraints. A lack of validity could also lead to limited transferability.

**Delimitations.** I also considered delimitations. Based on recent studies illustrating the importance of having an evaluation process that accurately assesses teacher performance, such as Heneman and Milanowski’s (2001) discussion about using multiple pieces of data to accurately assess teacher performance, I wondered if the tenure process could be counterproductive to the goals of the teacher evaluation system. I noticed that one of the largest problems in education today is that teachers receive tenure and that the school unions fight hard for teachers to maintain tenure status in kindergarten through twelfth grade.

Even though I questioned how tenure could impact the effectiveness of the evaluation process, readers need to understand that the data and analyses contain honest reflections of the research goals. Open-ended questions were written and utilized to give interview participants ample chances to share opinions and reflections. Middle school teachers as well as other doctoral students were engaged in the process to read the questions before they were used in interviews, in order to provide peer feedback and improve the clarity of the questions.
Then, each participant was interviewed for an extended period of time, with open-ended questions to ensure open, honest communication. Finally, after the interviews were transcribed, emergent methods were used to reconfirm that the data accurately reflected the perspectives of the participants (Hesse-Biber & Leavy, 2006). Each person interviewed was later sent documents, including the analysis of their interviews, to review to ensure they accurately reflected that teacher’s viewpoint. If the teacher felt it did not, the document was adjusted accordingly. This process allowed me to represent the interviewees more accurately and to be mindful about my own biases.

Diversification of the sample was also a key focus of the project. It was important to diversify the sample by grade level, school ranking, years of teaching experience, and geographic location. By diversifying the sample, more viewpoints were available for consideration when it was time to analyze the teacher evaluation process.

By reading *Qualitative Research Design* (Maxwell, 2013) and *Qualitative Inquiry and Research Design* (Creswell, 2013), additional strategies were identified for helping to minimize the impact of researcher biases and research threats when analyzing the data toward drawing sound conclusions. Strategies for increasing validity include engaging in peer review and providing rich descriptions (Emerson, Fretz, & Shaw, 2011). By engaging in peer review, the researcher receives an external perspective to keep the researcher honest and on task. Peers ask difficult and clarifying questions about meaning and methods. By engaging peers, information can be viewed through a different lens, which helps increase validity. Rich, detailed descriptions provide a reader with thorough analysis and conclusions. From the description, the reader makes their own decisions about how to use and apply the information. During this study, rich descriptions were used for writing field notes, analyzing
data, and constructing concluding statements. These additional strategies were reasonable to conduct and cost and time-effective.

**Participant selection.** In order to ensure quality data collection and analysis, I selected participants purposefully. Purposive sampling focuses on “the assumption that the investigator wants to discover, understand, and gain insight and therefore must select a sample from which the most can be learned” (Merriam, 2009, p. 77). Purposive sampling leads to in-depth studies and descriptive information. Knowing that teachers have large workloads and other time constraints, gathering a representative sample of participants was a primary focus. Though certain teachers were more assessable, the sample included teachers who work at a variety of schools. Observing and interviewing professionals at all levels of the teacher evaluation process and located in schools with various report card ratings is also important to developing a holistic understanding of the evaluation process. Without a purposive sample, researchers tend to craft inferences that are not well-grounded, nor generalizable (Miles, Huberman & Saldana, 2014). “The real problem with selective sampling and hasty generalizing is that you can slide incrementally into these biases, with the first layer preparing the ground for the next” (p. 295).

Polkinghorne (1989) recommended interviewing a minimum of five participants who experienced the phenomenon. He also stated that no more than 25 participants should be interviewed. To stay within the recommended range, five participants were selected from public schools. Each co-researcher met six important criteria. All participants: a) currently taught in elementary or middle schools, b) worked in a public school district in the southwestern portion of the United States, c) had been involved in at least two teacher evaluation cycles, d) have taught for over five years, e) had experienced the phenomenon,
and f) were interested in reflecting on their experiences with the teacher evaluation system. Each participant experience involved a “lengthy person-to-person interview that focuses on a bracketed topic and question” and possibly follow-up interviews (Moustakas, 1994, p. 104).

At the onset of this study, the southwestern community included in the research used a state-wide school ranking system wherein each school was assigned a school grade (A-F). However, at the time of the final narrative, the state had decided to discontinue the A-F ranking system. When the A-F grading system was in place, all schools received points for various components, and the points were added together to assign the school grade. In this southwestern state, elementary and middle school scores are competency-based, while high school scores focus more on graduation rates and preparation for college and jobs. In this study, all school grades, A-F, were represented in the sample. Maximum variation, this type of sampling, allows the researcher to establish criteria based on characteristics to maximize differences in order to receive various perspectives (Creswell, 2013). Maximum variation allows researchers to be involved in “identifying and seeking out those who represent the widest possible range of characteristics of interest in the study” (Merriam, 2009, p. 79). Marshall and Rossman (2010) stated that researchers need to remain flexible during their research, and samples may change. To achieve the appropriate level of flexibility, a variety of teachers were spoken to, at multiple schools within the district and at various schools with rankings spanning from A-F, about potential participation in the research.

**Context of the study.** When engaging in phenomenology, all participants must have experienced the phenomenon. This is referred to as a *criterion sample*. Since all teachers in this southwestern state experienced the teacher evaluation process, all fulfilled the criterion. The teacher evaluation system in this district measured five components towards teacher
effectiveness, which included classroom observations, planning, preparation, professionalism, student achievement, parent and student surveys, and teacher attendance (see Appendix C). Teachers received scores in each area. The school district compiled the scores in each area to create a summative score. Teachers also received a report explaining the details of each score. The distribution of points in all five areas was dependent on the availability of student achievement data over the previous three years. When collecting data during observations, evaluators used an observation rubric (see Appendix B). This rubric contained domains, strands, and elements, which helped the evaluators break down the key elements observed during their classroom visits. The observation rubric focused specifically on ranking teachers in two of the five categories. The two categories included classroom observations and planning, preparation, and professionalism. The score for student achievement was based on student standardized tests. Parent and student survey data was compiled and assigned a score. Teacher attendance scores were based on the number of teacher absences during each school year. Once all five categories received a score and a summative score was calculated, the teachers received one of the following rankings: ineffective, minimally effective, effective, highly effective, and exemplary (Educator Evaluation, 2019).

Scores varied based on whether or not a teacher’s cohort of students had achievement data over a three-year period. If this data did not exist for the teacher, then student achievement counted as 0% of their overall score. In this case, classroom observation counted as 50%; planning, preparation, and professionalism was 40%; and the categories of parent and student surveys and teacher attendance both counted as 5% of the summative score. If a teacher had one to three years of student achievement data, that data counted as
35% of the summative score. Classroom observation counted as 40%. Planning, preparation, and professionalism was 15%. Parent and student survey data and teacher attendance both remained at 5% of the summative score (“What is NMTeach?” 2019).

For the classroom observation component, evaluators were required to complete a district level training focused on understanding the observation domains, timeline, and rubric before evaluating teachers. One completing the training, evaluators were also required to observe each classroom at least twice each school year. During observations, evaluators focused on the overall class culture, classroom management, student engagement, and student assessments. Student achievement focused on student progress over a three-year period on standardized tests. Various grade levels were required to take different tests each year. The category of planning, preparation, and professionalism was scored using submitted teacher artifacts, which were uploaded in a system for the evaluators to review. These artifacts included lesson plans and student work, with a focus on pedagogy, professional growth, understanding of the students and their individual needs, and overall communication with families. Parents completed surveys for students in kindergarten through second grade, while students in third grade through high school completed their own surveys. Teacher attendance was not negatively impacted by religious holidays, professional development, bereavement, or the Family and Medical Leave Act (FMLA). By including all five areas of measurement, the school district strived to create summative scores that reflect overall teacher effectiveness (Educator Evaluation, 2019; “What is NMTeach?” 2019).

Once participants were selected for the study, the field work began. Access was gained to schools where the researcher was acquainted with teachers and/or where family or friends attended school. Expanding the research to schools in less familiar areas of the
community became the focus. Access to those areas was achieved through personal or university connections or by contacting school administrators directly. Initial contact by phone was informal. First contacts with the teachers included an explanation of the research plan to gain teacher perspectives on the observation process within the teacher evaluation system. Each teacher was asked to participate in an interview, with open-ended questions (see Appendix A). It was explained to participants that interviews would provide the greatest opportunity to ensure their thoughts were accurately represented. To increase the validity of the interviews, emergent methods were utilized as a method of analysis. The emergent methods include listening for the plot and listening for contrapuntal voice (Hesse-Biber & Leavy, 2006). By using these methods, the process validity was improved by utilizing triangulation analysis to find overlapping patterns, categories, and themes. The outcome validity was also increased by asking participants to verify the use of these methods, and positively impacting the transferability of the research.

During the initial contact with the teachers, the presentation was focused on the research and introduction of the researcher, with the goal of communicating to the teachers that the project was teacher focused and how it was organized. Further goals included conveying an attitude of open-mindedness about what teachers might share, as well as working to form an impression of the situation as friendly, approachable, and honest. The hope, it was relayed, was to accurately understand the teachers’ perspectives of teacher evaluation process. Lastly, the assurance of confidentiality of both the feedback and data collected was provided, to make sure the teachers felt safe sharing their thoughts as a participant of the project. After this initial presentation, teachers could ask questions or provide comments.
Methods of data collection. The primary method used in this qualitative research project involved in-depth interviews with participants who have experienced the phenomenon (see Appendix A). Interviewing is a “conversation with a purpose” and used to gather “a special kind of information” (Dexter, 1970, p. 136; Merriam, 2009, p. 88). An interview helps the researcher discover what is “in and on someone’s mind” (Patton, 2002, p. 341). As a researcher, interviewing helps understand the participants’ perspectives. The data source generated from this method included transcripts.

Another method used in phenomenology is documents. The data source generated from this method is the document itself. The written documentation included emails between the researcher and participants used to engage in member-checking. During the interview process, teachers shared some of evaluative documents, especially as trust and rapport was developed. Member-checking, another common method used in phenomenology, is the process of reviewing interviews with participants to guarantee that accurate knowledge is emerging from the data. Lastly, the findings were compared with a researcher’s journal which was utilized as a data source, generated from analyzing an interpretation of the researcher’s own thinking during the process.

By using three methods and data sources, triangulation occurred and patterns arose, leading to categories of data. Categories of data are the “collection of similar data sorted into the same place, and this arrangement enables the researchers to identify and describe the characteristics of the category” (Morse, 2008, p. 727). A theme, however, is the “meaningful essence” found throughout the data (p. 727). Triangulation allowed for the review of patterns across all the data sources and methods, and the opportunity to engage in cross-checking (Merriam, 2009). These patterns led to data categories, which ultimately aided in the
development of themes. Triangulation also helped validate compiled data by the observation of patterns, categories, and themes. This process helped verify findings and decreased possible risks of biases.

Before each interview, there was time set aside to reflect on personal biases and to achieve epoché, which is the practice of “suspending one’s various beliefs of the reality of the real world in order to study the essential structures of the world” (van Manen, 1990, p. 175). After personal reflection, the researcher conducted semi-structured interviews, which took place outside of instructional time. During the interviews, an introductory script was read, a consent form was reviewed, and flexible questions were provided, which offered “a mix of more and less structured interview questions” (Merriam, 2009, p. 89). The consent form explained participation in the study and mentioned permission to audiotape the interviews, as audio recordings were made during interviews. These recordings were later transcribed, with any identifiers replaced with study codes. Interview questions included:

1. What has been your overall teacher evaluation experience throughout your career?
2. What has been your overall experience with the observation process within the teacher evaluation system?
3. What is the purpose of the current evaluation process, from your perspective?
4. In this southwestern state, how effective is the current observation instrument, from your perspective?
5. What components of the observation process are most useful, from your perspective?
These open-ended questions provided details about the participants’ experiences. Moustakas (1994) discussed the importance of using broad questions to “facilitate the obtaining of rich, vital, substantive descriptions of the co-researcher’s experience of the phenomenon” (p. 116).

I followed Moustakas’ (1994) guidelines for interviewing. Moustakas suggested that researchers develop open-ended questions prior to the interview, yet the questions are “varied, altered or not used at all when the co-researcher shares the full story of his or her experience of the bracketed question” (p. 114). The questions and answers remained open if participants decided to expand on a specific story. In addition, Moustakas mentioned starting the interview with a brief social conversation to help the co-researcher relax and feel comfortable and then asking the participant to take a minute to focus on the phenomenon before sharing their experiences.

During the interview, raw field notes were collected. Miles, Huberman and Saldana (2014) suggested taking raw field notes using personal abbreviations and sketches. “Field notes taken during an interview usually contain a fraction of the actual content” (p. 71). The purpose of the raw field notes was to write up more formal field notes following the interview. Abbreviations and sketches helped jog the memory for writing a more extensive description.

Commonalities were hunted for within the collected data. Transcripts were coded, as were documents, using descriptive codes and sub-codes. Patterns, categories, and themes emerged within the data. The conclusion of this research is transferable and generalizable due to purposeful sampling.

Methods of analysis. During the analysis, descriptive coding was used, which means that labels were assigned, such as EVAL, to summarize the type of data that reoccurred.
Sub-codes were also used. A sub-code is “a second-order tag assigned after a primary code to detail or enrich the entry” (Miles, Huberman, & Saldana, 2014, p. 80). Sub-coding allowed for an extended coding that provided additional detail. For example EVAL:PEER was used to indicate teachers who expressed interest in having a peer evaluate them as part of the evaluation system. When the coding process started, a starting list was used to developed and gather multiple codes while still working through the data. Coding helped to realize multiple patterns. Examples of patterns that emerged included relationships with their evaluators and the impact of those relationships on the observation feedback. In order for these patterns to emerge, most participants provided details within the areas mentioned. For example, when all the participants mentioned how evaluator relationships with the teacher impacted the observation data, then a pattern emerged, which needed further study. Data was reviewed at least three times to confirm that no important patterns went undetected.

After analyzing the various patterns multiple times, specific patterns were noticed to be reoccurring frequently across all data sources, so categories were developed with codes and sub-codes beneath each category, to provide additional detail. At this time, final categories were discovered. Table 1 shows examples of possible codes and sub-codes.
A description was provided for each code and sub-code. For example, EVAL:REL was noted as teachers describing how their relationships with evaluators made an impact on the feedback provided from the evaluation. EVAL:ACC indicated teachers questioning the accuracy of the observation process within the evaluation system. As the analysis continued, certain sub-codes were added and others were realized to be less valuable through the perspectives of the teachers, upon triangulating the data. Certain categories eroded as others took center-stage during continuous analysis. From the categories, methods and data sources were reviewed. Themes were discovered using triangulation and by analyzing what categories emerged most often and across all data sources. This analysis allowed the themes to emerge.

After finding patterns and categories, and creating themes, the data was condensed and reduced further. A matrix was constructed and ideas that were presented by every data source and by most teachers were noted. This activity helped condense the ideas within each theme even further. By engaging in the matrix activity, strong connections within each theme were realized.
Conclusion

To better understand teacher perspectives, phenomenology provided the opportunity to collect the perspectives of multiple teachers. Phenomenology led to a better understanding of what teachers thought about the evaluation system and how they believed it could improve. By working with teachers, commonalities were discovered, which led to the exploration of multiple concepts within the evaluation process. Because teachers consciously engage in the evaluation process, phenomenology allowed me to analyze the crafted experiences of each participant. By focusing on phenomenology, people’s true lived experiences became the center of the research. The focus remained primarily on the essence of experiences within the teacher evaluation system (Creswell, 2013). It was discovered, by studying five individuals through interviews and data analysis, that certain phenomenon existed in regards to teachers’ perspectives towards the observation portion of the evaluation system. Bracketing occurred in order to separate the researcher’s own experiences from those of the study participants, in order to better understand their perspectives. Bracketing helped maintain a blank slate when collecting and analyzing data. Transcendental-phenomenological reduction allowed a focus on the essence of the experience and helped to better understand the original structure. This allowed separation of the phenomenon. By concentrating on epoché and transcendental-phenomenological reduction, understanding positionality, focusing on participant selection, using various methods for data collection, and analyzing data thoroughly, a phenomenological study was conducted that ensured quality data collection and analysis.
Chapter 4: Catie

Overall Teaching Evaluation Experience

Catie is currently a head special education teacher at a C-ranked elementary school in an urban public school located in the Southwest. Catie has taught for twenty-two years. Until August 2018, Catie taught in the same school for her entire career. She taught kindergarten all the way to fifth grade in every setting, from full-inclusion to self-contained. Catie has worked with over six evaluators and has been evaluated every year. She always received high evaluation scores until her most recent principal moved to her school. When the district introduced the new evaluation system, Catie was encouraged that the district evaluation appeared to reflect the Danielson Model (Danielson, 1996). In theory, she believed that a strong evaluation system encourages teachers to share ideas with each other, improves teacher practices, and helps areas of improvement to surface. However, Catie found the reality of this evaluation very different than expected, and stated, “This supposedly objective system of evaluation is completely subjective.” After receiving highly effective and exemplary scores in the first few years of the new evaluation system, Catie moved to minimally effective when her administrator changed and her students started taking state standardized tests. Ironically, during those school years, Catie taught the same group of students. She taught them in kindergarten and in second, third, fourth and fifth grades. She followed them throughout the various grade levels due to her special education training and the specific needs of the students. However, when the students entered third grade, they started taking the state standardized tests. Because a number of her students fall on the autism spectrum and have learning disabilities, they often can’t read the state standardized tests. Even though the state standardized test has speech software, “autistic children do not
Catie spent hours explaining to her students various strategies for taking the tests.

I mean we spend everyday teaching them how to highlight information that is important for a math problem and teaching how to draw a picture. If they don’t know what to do, think about what you know and draw a picture. These are all common sense things, but this is difficult for a child who has a processing disability or a language disability. This is where the test really changes the game.

Catie was frustrated that her students would never be able to pass the test, and her evaluations would continue to suffer. While her students continued to struggle with the state standardized tests, Catie recognized that it was almost impossible for her overall evaluation scores to become effective again, even though her students composed data notebooks that illustrated their progress throughout the school year.

I have a kid who I joke is the only kid who can fall off the floor, because he is so ADHD along with autism, and now he is sitting in a general lab for language arts. Is he going to do well on [the test]? No, but he has shown progress, but they don’t care about that. All they care about is that he is still under grade level.

Catie stated that the focus should be placed on showing progress, not just on achieving grade-level standards. In her opinion, her evaluation scores should reflect her ability to help students make strides towards achieving their IEP (Individualized Education Program) goals.

**Observation Process within the Teacher Evaluation**

When asked about her experience with the observation process within the evaluation system, Catie mentioned that at the beginning of the school year, the evaluator asked all teachers to sign up for two formal observation dates. Catie tended to “drag her feet.”
Because she did not sign up, the evaluator sent many emails to solidify two dates. “I tell her [the evaluator] show up whenever you want; I am not going to put on a show for you.”

The evaluators asked for lesson plans before each formal observation. However, Catie sent them a “lesson structure” because her plans might change due to the performance the day before. She might end up having to reteach or change the plan altogether. Catie continuously assessed the needs of her students and adjusted lesson plans. On every occasion, her evaluator arrived late when completing a formal observation. Because Catie worked in a special needs classroom, she never knew what might happen during the observation. “Halfway through a lesson, it is possible I might have to stop and change a diaper on a fourth grade boy.” Her lesson plans always look different than lesson plans in a general education classroom. Sometimes her evaluators understood the needs of her classroom and sometimes they “have no clue what a special ed program or classroom looks like.” The inconsistency was challenging for Catie. The district asked Catie to videotape SPIRE (Specialized Program Individualizing Reading Excellence) lessons, so they could show and educate other principals, yet her own principal gave her lower scores in this area on her evaluation. Also, depending what happened in the classroom earlier in the day, Catie might have to change her lesson plans all together. The evaluation system makes these types of changes challenging. The evaluator expected to see one thing, and then things completely changed.

Catie worked with four evaluators over the last few years. She found that their experience in understanding the special education population really impacted her scores within the evaluation system. Her first evaluator never provided much guidance or feedback because he tried “to make everyone happy.” She found her second evaluator more helpful
because “he knew he didn’t know what he didn’t know and would ask.” He asked her why she did certain things during her lessons. Normally, she explained the various IEP goals and how she incorporated those needed skills into the daily lessons. Her third evaluator “asked us to do all kinds of things that were unethical and illegal out of ignorance.” He left the school after one year, and she did not find him helpful. Her fourth evaluator was amazing because she truly understood the challenges of special education and supported the teachers. The evaluators greatly impacted the effectiveness of the evaluation process.

Catie set up pre- and post-conferences with her evaluators. Catie stated that the evaluation was subjective, so she appreciated having conferences to explain her lesson plans, clarify aspects of her lessons, and answer evaluator questions. During the pre-conference, she discussed the parts of her lesson and the reasons for the layout. The evaluator and Catie discussed the lesson plan. The post-conference was not as much of a discussion. During the post-conference, the evaluator explained the scores and reasons behind each score. If she disagreed and did not understand the reasoning behind a score, Catie pushed back and questioned the evaluator. Occasionally, the evaluator adjusted a score based on the clarifying information proposed by the teacher, but an adjustment was extremely rare.

The evaluator also came to Catie’s classroom for one walk-through during the school year. “I think that one has more to do with environment, engagement of kids, and organization of your environment. Are the kids all engaged? What is your rapport with the kids? That kind of thing.” The evaluator provided Catie with a follow-up email indicating what she observed during the short visit.
**Purpose of the Teacher Evaluation**

Catie believes that the purpose of the evaluation is to break apart the education system. The evaluation system to her “feels punitive.” Catie knew principals who told their faculty that no one at the school would receive a score of four or five.

So, if I can show you a five, you can’t sit there and say, before we’ve even met, before you’ve even observed me, that I can’t have a five, but I know for a fact that our previous administrator got taken to task for having too many people in the *highly effective* column.

Catie stated that if an evaluator documents why a teacher earned high scores, the teacher should receive those scores. The evaluation should be used to help teachers improve, not to discourage them.

**Effectiveness of the Observation Instrument**

Catie does not think the observation instrument is effective because it does not help teachers grow professionally. In her opinion, the focus should be placed on student progress, not on student test scores. As mentioned prior, some of her students will never achieve grade level status on the test, leading her evaluations to continue to be negatively impacted. “I got 3’s and 4’s on everything, my previous evaluator – 3’s, 4’s, and 5s --, but this one doesn’t give me 2’s, but it’s that piece of how you add in that value-added score from testing, and I get tanked.” Catie had to basically receive perfect scores in all other areas for her evaluation to improve due to the large impact of the testing. As a result, Catie is no longer motivated by the evaluation system.

Catie originally felt excited about the new evaluation system, recognizing that sometimes “it takes outside eyes” to help you improve in the classroom. However, after
being involved in the system, she recognized that the principals seemed extremely “rushed”
and did not have the expertise to always understand what certain classrooms were supposed
to look like. Often, the evaluators came in late to observe a lesson. They missed key
components of the lesson, like the review of routines and procedures, which led the rest of
the lesson to seem confusing. Teachers got marked down for something the evaluator did not
see. Catie’s most recent evaluator did not ask follow-up questions when she was unclear
about strategy used during a lesson. Catie challenged her evaluator in a few areas, but she
never changed any of the scores. Sometimes, “they don’t know what you are doing because
you are teaching a specialized program, and this is where it goes crazy.” Effectiveness of the
instruments decreased when evaluators were too busy to spend adequate time observing
classrooms, and to truly understand all aspects of the lessons and when they did not
completely understand the components of certain lessons. The capabilities of the evaluators
really impacted the effectiveness of the observation instrument. Because Catie did not
receive authentic feedback from the evaluation, she reached out to the behavior support team
and asked them to observe her classroom and provide feedback.

The observation instrument is also not effective because the district pressures
administrators to avoid letting scores get too high. “I think she [the administrator] is under
pressure for your observation to have you be a little bit lower and then on your second
observation have a little bit of growth.” Catie did not think the scores on the evaluations
were accurate. And, when Catie collected data and documents to show she deserved higher
scores, the evaluator seemed “afraid to open the door of debate over these because then she
would have to go back and change things.” Catie spent hours uploading documents to
prepare for the observation and to help the evaluator understand the components of her
classroom, yet she still received *minimally effective*. Because the evaluator knew Catie felt upset by the results, the evaluator did not review the scores with her in person. “She doesn’t like to see me cry.” Catie believes that the pressures from the district office lead to inaccurate scores.

The effectiveness of the observation instruments also diminishes when human factors are not considered. Students have many challenging situations happening in their lives that impact their academic performance. When evaluators do not focus on the whole child, the authenticity of the data collected and the use of the observation instrument declines.

**Components of the Observation Process**

During the observation process, Catie engaged in a pre- and post-conference with her evaluator. Catie spent hours uploading documents to help the evaluator understand the specific needs of her classroom prior to the pre-conference. During the pre-conference meeting, the evaluator and Catie discussed the lesson to clarify the objectives. Catie thought this time was useful because it was important that the evaluator understood why Catie structured her lessons in a specific way. During the post-conference, Catie did not discuss the scores with her evaluator. She was basically told why the evaluator gave her the scores. When Catie questioned a score, the evaluator did not make changes. Catie found the time spent in the pre-conference to be more useful than the time spent in the post-conference, and Catie did not find the time spent uploading material and artifacts to be a useful component of the observation process.

Often, the evaluator showed up late on the observation day and missed part of what was discussed in the pre-conference. Her evaluator brought an iPad and completed a checklist. The faculty at Catie’s school had been asked to not share the scores on the
checklists with each other. By not sharing data, Catie mentioned that there is a lack of transparency and teachers did not receive the opportunity to collaborate and learn from each other. The data she received ranks her performance in domains one through four. When evaluating the domains, evaluators ranked teachers’ levels of performance as one of the following: ineffective, minimally effective, effective, highly effective, or exemplary. Her complete evaluation score encompassed scores from the four domains, which included test scores and personal absences. In the areas of domain one and three, Catie scored above the school, district, and state averages. When looking at domain two and four, Catie scored above her school average, slightly above the state average, and on par with the state average. “So, I am higher than other people in the state or equal to other people in the state, so why am I minimally effective?” Catie thinks it is unfair that test scores negatively impact her evaluation to such a large degree, especially considering the subgroup of students in her class. Catie did not find the calculation of her final score truly reflective of her overall effectiveness as a teacher.

**Data Collected during the Observation Process**

During the observation process, the evaluator collected information about the upcoming observation during the pre-conference and received the uploaded artifacts. Then, the evaluator used an iPad to complete the checklist, which focused on domains one to four. The evaluator might include a few other comments with the checklist. Catie indicated that the evaluator collected limited data. She thought that the data collected did not acknowledge the hard work she put into building an exceptional classroom. Then, during the post-conference, the evaluator told Catie her scores, with limited opportunity to collaborate or clarify aspects of the lesson. Catie believed that the accuracy of the data diminished because
her evaluators did not know what they were looking for or observing during a lesson. Sometimes, when they did not see something, they assumed it never happened in the classroom. Teachers need an opportunity to explain various aspects of their lessons and clarify evaluator questions to increase the accuracy of the scores. Her current evaluator did not provide time during the post-conference to clarify the lesson. When an evaluator completed a walk-through, Catie received an email informing her that the collected data had been placed in the system. The evaluator did not discuss the data collected during a walk-through directly with Catie. During this interview, when Catie talked about the accuracy of the evaluation and data collection in her post-conference, Catie cried because the process was so frustrating. Catie continuously tried to become a better teacher.

**Changes to the Observation Process**

When asked about needed changes in the observation process, Catie stated that the data collected did not accurately reflect her teaching. For example, when she received scores above average or on average for her observations, how did she end up receiving a *minimally effective* rating? She would never be able to raise her evaluation score when working with her current group of students because they continued to struggle on the state standardized tests due to their learning challenges. The test negatively impacted Catie and her students. Catie has seen her own students punch themselves and slam or throw computer keyboards due to the stress of the test. Some of her students could not read the test.

It does hurt children to take the test when it is not appropriate for them because we are saying the same thing we say to teachers. Look at everything you don’t know, and that is hard for a kid. As an adult, it is hard for us to hear.
In Catie’s opinion, the test should not count so heavily when calculating her final evaluation scores. Catie taught a student who has an IQ of 51 and is autistic.

I said to myself ‘I bet he is really hard to test, and even if the circumstances were great and all the stars and moon lined up and everything was perfect and he was really compliant, he was not going to jump 30 points to be in low average.

The test should be used differently for general education students than for special education students. The data needs to better reflect teacher performance, and test scores need to be used more fairly.

Not only did the test scores decrease Catie’s overall evaluation, but her new principal also did not understand exactly what her classroom was supposed to look like, nor did they understand the programs used in her classroom to support her special education students. The lack of expertise of her evaluator also negatively impacted her scores and made the data less accurate. Catie indicated that she had “issues” with her current evaluator. When Catie asked her evaluator questions about the observation scores, the evaluator got “very defensive.” The evaluator is an important part of the effectiveness of the entire evaluation program. Catie stated that the way evaluators are assigned and trained needs to change.

Catie understood that evaluators are pressured to show improvement, so she received a 2 in an area, and then the evaluator would raise it to a 3 after the next observation with little reason for the change. One of Catie’s prior principals told her “we are told to not rank everybody very high because you can’t show growth and you look bad.” The pressures placed on the evaluators also diminished the accuracy of the scores. Catie stated that the collection of data needs to become less biased and subjective in order to increase the effectiveness of the overall evaluation system.
Final Thoughts

When asked if she had any final thoughts, Catie immediately mentioned teacher absences. Teachers receive a benefit of 10 excused absences a year. If a teacher took more than one absence, they got penalized on their evaluations. Catie had more than two months of sick time, and last year she was absent for eight days due to significant health issues. Her absences contributed to lowering her evaluation scores. Her students reacted strongly when Catie was absent. As a result, she never took time off unless she was extremely ill. Catie revealed that she earned those eight sick days, and her evaluation should not be negatively impacted by a benefit.

Catie also mentioned the impact of surveys on her evaluation score. The district asked students and parents to complete surveys online. When her students took the survey, often they did not understand the questions. All teachers had been told not to reword the questions for the students. Catie stated that a question on the student survey might ask about a teacher knowing when her students are upset, and the student might think that was asking if they upset their teacher. Misinterpretation of questions could lead to dramatically different survey results. Also, when the district asked parents to complete surveys, only about half of Catie’s parents complied. This led to incomplete data. Some of the parents did not know how to read, which added an additional challenge. Because the absences, surveys, and test scores count as a large portion of the evaluation, Catie will never be able to achieve a score higher than effective. Catie stated, “Why am I even trying?” However, while Catie will never stop improving and putting forth effort to help her students, she will “stop putting on a performance or a show for somebody who doesn’t know my kids, who doesn’t know why I am doing what I am doing.”
Chapter 5: Debbie

Overall Teaching Evaluation Experience

Debbie teaches science and Spanish in a B-ranked public middle school located in the Southwest. She has been teaching for ten years. Debbie served as a student teacher in the southwestern public school district. Then, she taught in a private school for one year and decided to return to the public school district. She has worked at the same middle school for nine years. Debbie worked with four different evaluators in the last six years who observed her class once or twice a year. Each visit lasted 30 minutes at the most, which is less than a full class period. Even though Debbie received high marks on her evaluations, she questioned the accuracy of the data collected during observations. “Is it helping me any?” She questioned the benefits of the evaluation program and did not think it pushed her to grow as a professional.

Debbie noticed that the evaluation process seemed different depending on the evaluator. A few years ago, Debbie’s evaluator popped into her classroom unexpectedly and indicated that she was completing her observation. Debbie did not meet with the evaluator before or after the observation. Debbie received scores online. Another year, one of Debbie’s close friends evaluated her performance. “She didn’t even come to my classroom, and she says ‘I know you do a good job,’ so she gave me all 4’s and 5’s.” Her friend never came to see her teach, yet Debbie knew that her friend observed other teachers’ classrooms. Debbie thought that there was a bias in the collection of the data. This past year, her evaluator walked through the classroom for a few short visits and completed one formal observation, even though two observations were planned. The evaluator met with Debbie before and after the observation. She provided Debbie with feedback. “At least, I thought,
well there are some things I can work on, and she did give me some resources, which was nice.” Even though Debbie appreciated the feedback, in her opinion the evaluator did not fully grasp what was truly occurring in the classroom during the walk-throughs, and the evaluator only visited her classroom for 20 minutes during the formal observation. Because Debbie received either minimal or no feedback at all, she found that asking the students for feedback throughout the school year was the most valuable for growing professionally. Debbie found student feedback to be “brutally honest,” which she appreciated.

**Observation Process within the Teacher Evaluation**

When asked specifically about the observation process, Debbie explained that evaluators plan to formally observe each teacher twice during the school year. Debbie scheduled one of the observations, and the other observation was unplanned. During the planned observation, the teachers could “showcase” their best lessons. She thought that “you can kind of play the system a little bit if you are smart about it,” which undermined the accuracy of the collected data. Even though each teacher knew to expect an unplanned observation, this observation tended to increase teacher stress. Regardless of the type of observation, Debbie stated that the evaluators “are not getting an accurate representation of what it is like in class.” When the students noticed a visitor, they acted differently. It changed the dynamics of the classroom.

During Debbie’s observations, some evaluators walked around the classroom to look on the walls and talk to students, and others just sat in a corner. The type of data collected during the visits depended on the strategies used by the evaluators. One of Debbie’s evaluators wrote a minute-by-minute data collection, while a different evaluator provided
general statements about student activity and teacher responses. This year, Debbie met for both a pre- and post-conference with her evaluator, which did not occur in past years.

**Purpose of the Teacher Evaluation**

Debbie believes that the purpose of the teacher evaluation system is to “evaluate classroom environments” and analyze the academic achievement of students through the use of the teachers’ strategies. “What are the students doing and what is the teacher doing?” The evaluation helps teachers grow and reflect on their practices to help students learn. Debbie does not think that the current system helps teachers grow and does not align with the purpose. Without accurate data, how can teachers receive accurate scores? If evaluators do not accurately understand what happens in teachers’ classrooms, they can’t provide accurate data, which decreases the authenticity of the entire evaluation process.

Debbie thought that the purpose of the teacher evaluation was “to get rid of teachers who are not hip to the latest and greatest.” If teachers were not flexible and willing to make changes to their practices, the district wanted them to leave. Debbie mentioned that pushing teachers out of the system probably helped save money. When evaluators started questioning strategies used by veteran teachers, often times, the veteran teachers did not feel appreciated or respected. As veteran teachers decided to leave, the district received the opportunity to hire less-experienced teachers at lower salaries. “Honestly, I think it is a money game because it is a really good way to get inexperienced teachers who are willing to do anything to keep their jobs.” Debbie knew teachers at her current school that left because of the new evaluation system.
Effectiveness of the Observation Instrument

If the primary purpose of the evaluation is to help teachers improve and increase student achievement, Debbie thought the current observation instrument was ineffective. “There is no follow-through with these evaluations.” In Debbie’s opinion, evaluators need to provide opportunities for teachers to enhance their skillsets. An example includes providing opportunities to observe another teacher at their current or another school. Teachers receive their scores and that tends to end the conversation. Teachers and evaluators need to discuss ideas for increasing evaluation scores. How can we use the evaluation feedback to help teachers grow in the future? If a teacher received high scores, then Debbie thought that the evaluator believed everything was fine and only provided limited feedback. Debbie wanted to continue to grow and receive thoughtful feedback. If a teacher received low scores, the evaluator placed them on a performance plan. In Debbie’s opinion, the evaluators need to clearly state the objectives of the observation instruments.

Components of the Observation Process

Debbie found value in a few components of the observation process, including building a lesson plan and reflecting on how to construct it. She tried to integrate these practices into her everyday teaching.

As a beginning teacher, you are taught reflect, reflect, reflect, and look at, but as you start going on in the years, you forget to do that kind of thing and you forget that it is useful, and it is a good thing. So, it has been really neat to have that and to kind of make it mandatory so now it is like I have to do this.

Debbie appreciated being pushed to reflect on her practices. All teachers should continuously reflect to make sure they build lesson plans that align with the students’ needs.
Debbie found conferencing with the evaluator and collaborating with her peers useful. Before Debbie engaged in the observation process, she collaborated with her peers and shared her ideas about her upcoming lesson. She asked her peers for feedback to improve it. “We are able to sit down and rehash what are you doing in class, how is that working, and that part is really nice because we are getting ready for the evaluation, and we are helping each other.” Collaborating with her peers and her evaluator also helped Debbie reflect and improve.

Debbie did not find the observation data useful. Debbie believed the data would be more valuable if she worked with more than one evaluator. “Honestly, just having one person evaluate you, I don’t find that useful at all, because everyone comes in with their own ideas.” If a team of evaluators observed every teacher’s classroom, the accuracy of the collected data would improve.

Overall, Debbie did not find most of the observation components useful. She believes that her teaching had not improved or changed with the implementation of the new components. “It doesn’t feel authentic anymore.” Because of her desire to continuously improve, Debbie reached out to her students for feedback and used their feedback to adjust her practices.

**Data Collected during the Observation Process**

When asked about the data collected during the observation process, Debbie explained the evaluators used rubrics to compose scores for each domain. During the observations, the evaluators looked for evidence for each criterion. “It is really hard for an administrator to come in one time or maybe two times a year and to see if you are hitting all of the points on the rubric, because there are some days you do not hit them, and some you
don’t hit every single day.” Because of the limited amount of visits, the evaluators unrealistically expected to see all components of the rubric within a short period of time. Teachers either got points deducted for missing components or had to defend themselves once they received their scores.

The data collected during the observations depended on the strategies used by the evaluators. The data collected for Debbie over the past few years varied greatly. For the first time, Debbie’s evaluator asked to meet for a pre- and post-conference last school year. During the pre-conference, the evaluator asked Debbie, “What would you like me to watch?” As the evaluation only lasted for 20 minutes, the evaluator took notes on her computer and posted the comments in the computer system for Debbie to review. Then, they met to discuss the results. When working with a different evaluator, Debbie received scores even though the evaluator never visited her classroom. The evaluator, one of Debbie’s friends, indicated that she “trusts” Debbie was doing a great job and didn’t need to visit her classroom. A third evaluator “popped in” one time for a short observation. Debbie and the evaluator never met for a pre- or post-conference. The typed data showed up in the computer system for Debbie to review. These different observation experiences led the data to largely vary each year.

**Changes to the Observation Process**

Debbie stated that a team of evaluators need to observe each teacher’s classroom in order to collect authentic data. By having only one evaluator, the data tended to appear biased. In Debbie’s opinion, teachers and administrators should function as a team of evaluators. The teachers on the team should include someone from the same department and a previous mentor. By working with a team of evaluators, teachers would receive multiple perspectives.
When working with evaluators, the process needed to be more consistent. All evaluators needed to set up pre- and post-conferences with their faculty. Teachers desired concrete and useful feedback that encouraged growth. Debbie truly appreciated that this year’s evaluator gave her ideas and resources after observing her classroom. All evaluators should provide helpful resources and ideas for improvement instead of just completing a checklist and providing scores.

In order for an evaluator to understand the dynamics of a classroom and the needs of the teacher, evaluators need to visit classrooms more frequently. Debbie recommended that the evaluators visit her classroom for an entire week or unit. In order for evaluators to provide helpful feedback, they need to be present more often.

**Final Thoughts**

In Debbie’s opinion, receiving a score of 5 on an evaluation seems impossible. Evaluators tell teachers to be satisfied with 3’s. Teachers always tell their students to strive for high scores, yet the district gave teachers a very different message. “They [the evaluators] are like ‘no, just be satisfied with a 3; that is excellent; you are fine’ and not on a plan.” Debbie shared that the current evaluation does not motivate teachers to improve. Why would they strive for a higher score? Debbie mentioned that other industries provide incentives. She stated that providing teachers with incentives, such as receiving additional personal leave, time for professional development, or money, would motivate teachers to engage more in the evaluation process. The incentives should be “something that really makes you want to continue pursuing that extra step in going further and not just being satisfied with how you are teaching and how you are doing things every day.” When reflecting on the current model, Debbie did not think that anything was going to change
unless teachers received incentives or something else that helps motivate them enough to engage in the process. Right now, the evaluation “is just another hoop to jump through.”
Chapter 6: Kaitlin

Overall Teacher Evaluation Experience

Kaitlin teaches second grade at an A-ranked elementary school in an urban school district located in the Southwest. She has been teaching for nine years. In 2001, she taught third grade in a different district and decided that teaching was not for her. At that point, she decided to stay home with her young children. Then in 2012, after a divorce, she needed to return to work. At the point in her life, Kaitlin decided to return to teaching. She has been teaching second grade at the same school for the past eight years. Kaitlin’s experience with the teaching evaluation has been positive. She has worked with five different evaluators during her teaching career. As she realized that many teachers felt “hostile” towards the changes with the evaluation system, Kaitlin had “never known anything different, so coming into education that was part of the deal.” During her time teaching, she had seen only little changes made in the evaluation system. The union communicated some needed changes to improve accuracy, and Kaitlin believed those changes were made. “It is a work in progress just like anything we do; Even in our own work you have to start somewhere and take a look at all of it.” Her overall experience had been positive.

Observation Process within the Teacher Evaluation

As someone who appreciates feedback and hopes to continuously improve, Kaitlin embraced the observation process. She enjoyed people visiting her classroom and wanted the evaluator’s feedback. Kaitlin recognized that teachers can always improve their practice, and the observation process allowed teachers to receive feedback to reflect upon to help them grow. “I always want to make my practice better, so if somebody can walk in and tell how to do that, I am all for it.”
Kaitlin signed up for her official observations at the beginning of the school year and indicated if she planned to teach math or language arts. The evaluators liked to see one lesson of each curriculum area. During the observation process, the assistant principal visited Kaitlin’s class. While she observed for over an hour, the assistant principal walked around the classroom, looked at materials hanging on the walls, asked Kaitlin questions, and typed on her computer. “You feel like she set this time just for you.” The assistant principal tended to finish collecting data and writing about the observation during her time spent in the classroom. The assistant principal had consistently observed Kaitlin over the past few years. Kaitlin believed that having a consistent evaluator had advantages. The assistant principal witnessed implemented changes and growth in Kaitlin’s classroom. The consistency allowed Kaitlin and the evaluator to build a strong relationship, but Kaitlin also noticed benefits to receiving another evaluator’s perspective on her teaching strategies.

**Purpose of the Teacher Evaluation**

From Kaitlin’s perspective, the purpose of the teacher evaluation system was to help teacher’s improve. If teachers struggled with effectiveness, Kaitlin hoped that the evaluation process provided opportunities for them to reflect and grow. “I would like to believe that we do it so we can help struggling teachers become better, and we can get them the help that they need.” Though Kaitlin hoped that the evaluation would help people grow and/or steer ineffective teachers out of the profession, she expressed concerns that effective teachers were also leaving because they felt unsupported by the current system.

What it is doing is driving out good people, and people are retiring early. That is probably one of my biggest concerns about it. I feel like all of these unintended consequences that we won’t see until years down the road…You know where we are
spending money to replace these educators who were good and now we have less
effective people there.

Kaitlin believed the purpose of the evaluation was to help people grow, but she worried that the new evaluation system encouraged effective teachers to leave the profession.

**Effectiveness of the Observation Instrument**

Kaitlin did not think the observation instrument was effective. In order for evaluators to truly understand the dynamics of the classroom, they must be present in teachers’ classrooms more often. At her school, evaluators provided data from unannounced walk-throughs. They completed walk-throughs about two to three times per school year. During a walk-through, the evaluator looked for a few specific items and left after a few minutes. They completed a checklist, which included looking at classroom management strategies, posting essential questions and objectives, and more. Kaitlin encouraged evaluators to come for more mini-visits and to stay even longer. “Anyone can pull anything off for five minutes, but if you stay for 15 minutes, you might see a whole different process.” She mentioned that, by increasing the amount of visits and the length of each visit, the effectiveness of the observation instrument would improve.

**Components of the Observation Process**

When asked about the usefulness of the components of the observation process, Kaitlin indicated that having an evaluator that provides useful feedback is incredibly helpful. Because Kaitlin consistently worked with the assistant principal during the observation process, she had built a strong relationship with her and felt comfortable discussing all aspects of her classroom. The assistant principal had observed growth in Kaitlin’s classroom and worked with her closely to encourage additional progress. However, because Kaitlin
received scores of *highly effective*, sometimes Kaitlin thought that the assistant principal “only half listens because she knows I am strong.” Kaitlin recognized that benefits of working closely with the same evaluator, however she also hoped to work with another evaluator to receive a new perspective on her classroom. Kaitlin focused on continuously progressing and growing in her profession and said that, in order to continue to move forward, she needed a different evaluator in upcoming years. She hoped to work with the new principal to receive a different set of feedback.

   I feel like the feedback is useful when it is meaningful and thoughtful and applies to your lesson, and I do not feel like with her it is just cut and paste, and I have had that before with other assistant principals that we have had. It was just sort of cut and paste, and there is no value in it for me. I am not going to change. I am not going to grow.

As consistency with the evaluator helped with understanding the dynamics of the classroom and provided the ability for the evaluator to see growth over time, Kaitlin stated that having multiple perspectives to analyze her practices would be helpful. When the same evaluator observed numerous times, sometimes that feedback stayed the same or seemed very similar.

   As Kaitlin believed that every section of the evaluation serves some purpose, she questioned the usefulness of writing the lesson plan prior to the observation. She stated that anyone could write an effective lesson plan and “sound good.” However, does the lesson plan reflect the daily practices in the classroom? Kaitlin specifically pointed out the importance of incorporating domain four. This domain allowed teachers to describe “all those extra things that most administrators aren’t even aware.” Teachers’ jobs incorporate so
many tasks. To fully understand the effectiveness of a teacher’s performance, an evaluator must understand all the facets of their job, not just focus on the lesson plan.

**Data Collected during the Observation Process**

Kaitlin’s current evaluator took notes during her visits. She completed the entire collection of data while observing in the classroom. After walk-throughs, the evaluator uploaded a quick checklist into mylearningplan.com for Kaitlin’s review. Kaitlin did not find the information provided after walk-throughs useful. After the longer observations, the evaluator uploaded scores and a few additional notes. Kaitlin no longer met for any conferences with her evaluator. She received only written feedback, unless she disagreed with the assessments and scheduled an appointment. “It is kind of like we don’t have time for this anymore, so if you are okay with everything, just hit ‘submit’ is sort of where we have gotten now.” Not having conferences any longer negatively impacted the entire community. Conferences only occurred when a teacher really struggled in an area. Kaitlin no longer talked to any evaluator about anything. Kaitlin appreciated the feedback and data collected in the past, yet she noticed repetition in her feedback and needed a new evaluator to provide a different perspective.

**Changes to the Observation Process**

Kaitlin stated that all teachers should be engaged in pre- and post-conferences with their evaluators. At Kaitlin’s school, evaluators only conferenced with teachers when concerns about their practices surfaced. In order to explain all aspects of the lessons and learn from feedback, conferencing should be a key component to the evaluation system for all teachers. Teachers need the opportunity to provide context to the observed lesson. “It might not make sense to you because they have already had this background knowledge, or
we are finishing up a lesson, or we are doing a review.” By conferencing, all teachers would receive the opportunity to dig deeper into the feedback of the evaluator. Right now, the written feedback is brief and seemed “hurried and rushed to get on paper.”

In years past, Kaitlin met with the evaluator to discuss the observation. A few times, the evaluator made a few assumptions that were incorrect. The conferences provided Kaitlin with the opportunity to clear up the details. For example, Kaitlin had points deducted because the evaluator assumed that she did not work collaboratively with her students to build the class rules. However, Kaitlin had worked with the students and, during the conference, explained the process to the evaluator. Conferencing encouraged strong relationships between the teachers and evaluators and helped increase the accuracy of the collected data.

Kaitlin also mentioned the need for evaluators to visit her classroom more often. “I wish she would spend more time walking around and talking to my kids, asking them to see if they are internalizing what I am saying.” Kaitlin had great rapport with her students and thought that the evaluator would learn more about her rapport and other dynamics in the classroom if they visited more frequently.

Final Thoughts

A few years ago, as many teachers obtained high marks on their evaluations at Kaitlin’s school, the administration voiced concerns that PED (Public Education Department) might question the high scores given to the teachers. So, the administration stated that teachers were going to have to work really hard to receive a score of five. In order to receive a score of five in any area, the teachers needed substantial evidence showing teacher achievement at such a high level. Kaitlin admitted to being surprised that she received highly
effective as a new teacher. Kaitlin’s experiences with scores and comments by the administration brought into question the true accuracy of the data.

All teachers expect for evaluators to observe their classrooms. They just want to make sure it is fair and accurate. Kaitlin promoted more visits to her classroom. “It’s nice to have our administration come in and watch you do what you are so passionate about, to feel that with you and experience it.” Kaitlin loves teaching and strives to improve each year. As she did not like the feeling that she was getting compared to her peers through the evaluation system, Kaitlin understood the importance of receiving feedback and sharing her experience.
Chapter 7: Mary Gene

Overall Teaching Evaluation Experience

Mary Gene is an elementary special education teacher in an urban school district in the Southwest. She has been a teacher for 23 years and seven of those years have been in this urban school district. During her career, she taught kindergarteners all the way through fifth grade. She has worked with five different evaluators during her time in this urban district. Her school received an F ranking. Mary Gene has been evaluated for six years. She had received *highly effective* and *effective* marks within the current evaluation system. Through her experiences with the evaluation system, Mary Gene stated that the evaluation system did not present a clear picture in regard to teacher capabilities. Describing the process as “a dog and pony show,” Mary Gene mentioned that many aspects of her teaching were missed during minimal visits to her classroom. As she stated, the evaluation process is “not really a real clear picture of what you are.” Her experience with the teacher evaluation system had not been fulfilling nor helped her grow professionally. She stated that it is a time-consuming game that teachers have to play in order to prosper in their jobs. Mary Gene indicated that everyone knows which teachers should no longer be at the school due to their lack of professionalism, yet the evaluation process did not necessarily bring attention to the ineffective teachers.

Observation Process within the Teacher Evaluation

Mary Gene understood the importance of the observation process, yet she thought that the dynamics of the classroom changed when an evaluator was present. Teachers prepared lessons for the evaluators, which, often times, did not reflect what their true classroom environment the other days of the school year. For example, when teachers felt
nervous in the past about the potential actions of a student during an observation, Mary Gene had allowed teachers to send students to her classrooms while they were evaluated.

I mean if you are going to do it honestly, they should do a camera or something where they see reality because you are changing the structure. You may change it in a good way, you may change it in a bad way, but teachers are prepping for this one thing and then you know 364 days of the year they are doing something probably a whole different thing. I know it is necessary, but I don’t know how effective it is.

As Mary Gene shared that she had always been treated fairly during the evaluation process, she expressed concerns that the evaluation process did not accurately access the effectiveness of teachers.

Teachers’ evaluations are also impacted by students’ previous teachers and their effectiveness with the students in prior academic years. If a student has not learned information from teachers in the earlier years, the current teacher faces the task of helping that student achieve grade-level standards even though they might be multiple grade levels behind. If a student starts to get behind in early elementary grades, the gap gets larger each year, increasing the challenge for the current teacher to help them achieve grade level standards. The current evaluation focuses on grade level standards, not progress throughout the school year.

During Mary Gene’s experience as a special education teacher, she felt that the observation experience varied depending on the evaluator and the relationship between the evaluator and the teacher. She thought that administrators treated teachers differently depending on their personal relationships.
The administrators give certain teachers good kids because they like them and then new teachers, a lot of times, get loaded up with crappy kids. This is a new teacher. She will either sink or swim. They tend to do that a lot from what I have seen. The administrators have their favorites, so they give them all the kids that don’t have any behavior issues and the others, because I don’t like you, so give them the problem kids. So, there are all kinds of things that go into the whole thing that just don’t really show up in observations.

Mary Gene indicated that, in the past, she worked with administrators who greatly supported her and others that she felt supported her less.

While being observed, Mary Gene recognized a conflict between the daily expectations of the school and the primary areas focused on during the observation process. As she attempted to follow the curriculum provided by the school and district, Mary Gene prepared different lessons for her observations. The checklist used by evaluators focused on different areas than what she was encouraged to follow based on the curriculum. For example, Mary Gene encouraged students to lead learning opportunities in the classroom, yet there was little opportunity to provide time for student-led activities within the curriculum.

You give us curriculum, you give us science books, but it is not really about that.

You want us to do something on a day that is different, but everyday this is what we are doing and that is what you as a district are requiring. Why do you want us to do this to look good when in reality you tell us to do something else?

On a daily basis, Mary Gene used curriculum provided by the school district, yet during her observations, she planned “fancy lesson plans” such as student-led activities.
Purpose of the Teacher Evaluation

Mary Gene mentioned that the primary purpose of the teacher evaluation process is to show the challenges in the public schools, which might lead to school closures and reorganization. By using the state standardized tests, Mary Gene believed the public schools were put at a disadvantage when comparing them to charter and private schools. Charter schools used the state standardized tests less frequently than the public schools. Private schools used other tests to measure student achievement. Mary Gene wondered if the purpose of the teacher evaluation was to use testing to reorganize money within the school district.

Effectiveness of the Observation Instrument

When asked about the effectiveness of the current observation instruments, Mary Gene indicated that the tools at least promoted the idea of teachers thinking about the teaching process, which encouraged professional growth. However, in regard to students, the current instruments focused on grade-level achievement, not student growth. “I know kids who make huge growth. I have kids who have made two to three years of growth, but they aren’t on grade level.” Because evaluators used state standardized testing as the primary indicator for student achievement, Mary Gene mentioned that the observation instruments did not allow evaluators to observe other areas that illustrated student growth.

Nobody knows where the kids are except their teachers, so you know this kid is reading, doing great, but you can’t show anyone because nobody looks at that data. Nobody looks at DRA [Developmental Reading Assessment] scores or any other testing. It is just one test, the state test. That is all that counts.
If the observation instruments included an opportunity for evaluators to view other student assessments, evaluators would better understand the student growth occurring in each classroom. When the evaluators visited her classroom, Mary Gene stated that they received a limited viewpoint about the student growth that occurred each day. The “tiny snapshots” made it difficult for an evaluator to truly understand what happened in the classroom. If a teacher struggled during an observation and got “really rattled,” the teacher’s evaluation might not accurately reflect the teacher’s capabilities as a professional. The observation instruments provided only limited information. Mary Gene welcomed more frequent visits in order for evaluators to better understand the dynamics of her classroom and gain a more accurate picture.

**Components of the Observation Process**

From Mary Gene’s perspective, the most useful components of the observation process included conferences with evaluators and the combination of mini-visits with longer visits. Mary Gene desired post-conferences with her evaluator to review the observations, clarify any questions, and feel reassured “that things are going well and the kids are doing well.” She appreciated that her current evaluator prioritized ample time to spend with her and provided useful feedback even though he had a stressful schedule.

He is very good about this. He may see a problem, or say maybe you need to work on that, or I could support you in that. He has been good about having us go to different schools to observe different teachers, so that has been good especially for new teachers.

Mary Gene appreciated the feedback provided during the post-conferences. The feedback promoted self-reflection.
Mary Gene encouraged evaluators to visit her classroom as much as possible in order to better understand the dynamics of her classroom. Her current evaluator spent about an hour observing her students and watching Mary Gene teach. He also visited her classroom two to three times and spent a shorter amount of time. The combination of mini-visits and longer visits provided Mary Gene with the opportunity to check-in with her evaluator more frequently. It also helped the evaluator better understand the classroom dynamics.

Mary Gene did not find the paperwork associated with the observation process helpful. In Mary Gene’s perspective, many teachers, especially new teachers, felt overwhelmed by the paperwork and preparation that occurred before each observation. As Mary Gene stated, “It is a massive thing you have to go through.” Mary Gene believed that the current observation components did not lead teachers to receive the support and mentorship that they needed to improve.

**Data Collected during the Observation Process**

According to Mary Gene, during her observations, evaluators looked to see if she posted standards, viewed the material written on the board, and observed the interactions between her and her students. The evaluators completed a checklist while visiting the classroom. Mary Gene indicated that she prepared her classroom before each observation to reflect those standards listed on the evaluator’s checklist. “I have my standards up, I got this, I got that, no I’m not going to scream at you today.” Because teachers prepare their classrooms before observations, Mary Gene questioned the accuracy of the data collected during observations.
You’ve already threatened them to death so that they know if they act up, tomorrow they will pay for it. So, you are not really seeing how that teacher really talks to their kids. You are just kind of seeing a ‘dog and pony’ again.

Teachers use different strategies and techniques when being observed and “pretend this is what you were doing all along.” Mary Gene stated that teachers create lesson plans that cater to what evaluators want to see, and evaluators collect inaccurate data, because teachers teach differently when being observed.

**Changes to the Observation Process**

Mary Gene stated that the observation process should focus on student growth, not grade-level achievement. Teachers should receive evaluation data based on student growth and learning. In order to focus on student growth during the observation process, evaluators must know each individual student, their strengths, and their challenges. Evaluators should focus on observing how teachers address each individual student’s needs, including emotional and physical needs, not just academic. Evaluators also should consider the “human factors” associated with teachers and students. During an observation, a teacher or student might be having a “bad day.” If an evaluator is not visiting often, they might not realize that the specific day is an exception to the norm.

Some of the kids are going to be having a bad day, and some of the kids are sick and not going to be involved, but those things count against you. If a kid is not feeling well, I call their parent, lay them down in a room with a blanket, and let them rest a little bit, but if they came in and saw that, that would be really bad for me, but that is what that kid needs.
Mary Gene encouraged evaluators to visit classrooms more frequently in order to gain “an honest knowledge” of each classroom and understand the needs of the individual students.

**Final Thoughts**

When asked to share any final thoughts, Mary Gene expressed concerns about the incorporation of absences within the teacher evaluation process. Some years, teachers do not need to use their personal days, and other years might be more intense. For example, after not missing many days in prior years, Mary Gene had to be hospitalized for five days this past school year. Following her hospitalization, she fell and needed additional medical attention, leading to more absences from school. These absences negatively impacted her evaluation. “Every day counts against you, so if you have 1 day is 90%, 2 days is 80%, and you lose those points, and it will really kill you.” She also shared a friend’s experience with absences. Mary Gene’s friend delivered a baby, and within the same timeframe, her husband had a heart attack. Clearly, she needed to take time off of school. However, her friend experienced the negative impacts of absences on her evaluation. Mary Gene expressed that teachers should be given a few absences without penalizing the evaluation. After a set of days are used, then a penalty seems fair.

Mary Gene also expressed concerns that teachers did not play a large role in creating and implementing the teacher evaluation system. Teachers do not only teach students in the classroom, they also hold so many different roles to support the students. These roles are not always reflected in the evaluation system. “They are doing other things that are as important in the community, yet none of that shows on the evaluation.” From Mary Gene’s perspective, teachers need to focus on helping students become “better human beings.” The evaluation needs to incorporate other important aspects of teaching including building
rapport, teaching empathy, and discussing the meaning of success. “I want them to love
learning, and love each other and learn how to be good human beings more than anything,
because that makes you a smarter person all the way around.” If more teachers were
involved in the development of the evaluation, they could help the evaluation reflect all of
the aspects of teaching more accurately.
Chapter 8: Melinda

Overall Teaching Evaluation Experience

Melinda is a kindergarten teacher in an urban school district in the Southwest. She has taught kindergarten for 12 years in this urban school district. Melinda has worked with seven different evaluators during her teaching career. Her school is currently ranked D. Melinda shared that, before the implementation of the new evaluation system, the objectives of the prior system seemed very clear. With the new system, teachers feel stressed and concerned about the stability of their jobs. In her opinion, tying the evaluation to test scores was increasing teachers’ stress levels even further.

Throughout Melinda’s experience with the new evaluation system, her evaluations consistently contained inaccurate information. Her evaluation scores moved from highly effective to effective based on test scores. When she looked at the data more closely, Melinda noticed that the test scores reflected results for only three students. However, Melinda taught 21 students. During the prior school year, Melinda taught 21 students, and all students completed the test. The test data on her evaluation showed the results for 19 students. Out of the 19 students represented, five of the students were not in Melinda’s class. During a different school year, the evaluation reflected data for 12 students, when 15 students completed the assessment. As a reflection of the test scores, her evaluation slightly decreased over the past few years. “So, it is frustrating because all of these scores are added up for my evaluation when it is not accurate information that they are getting their data from.” Her evaluation also inaccurately stated that she achieved perfect attendance. However, due to a hysterectomy, Melinda went on leave for two months. She anticipated low scores due to her attendance. “We are not upset about being evaluated, we are always
evaluated, but we want to make sure it is a fair evaluation system.” At this time, the union asked teachers with inaccurate information on their evaluations to submit a description on the union website. The union built a case against the state. Melinda explained that, after the lawsuit between the union and education department, which incorporated the inaccurate data experienced by many teachers, the education department reduced the weight held by test scores. Melinda, disappointed by the results of the lawsuit, hoped for a “revamp of the whole evaluation.” When Melinda mentioned the inaccuracies of the evaluation to her principal, her principal recommended that Melinda not bring this information to PED’s attention. If they recalculated the data, depending on the test results, her scores could decline.

Melinda expressed concerns about how attendance, surveys, and test scores impact teachers’ evaluation scores. Because absences negatively impacted the evaluation score, teachers attended school sick. Teachers knew that the sick days would count against them, so they pushed through the day. How are teachers coming to school sick when we won’t allow students to come to school sick? Surveys also impacted evaluations, yet only about 50% of parents at Melinda’s school completed the surveys.

Fifty-percent of them are very active, very involved, and the other half are the ones you have to worry about checking folders, making snacks, helping with their kids at home, so I don’t think that is an accurate way to count in our evaluations. Melinda recognized that students would be tested on an annual basis, yet she stated that the weight of testing should be lowered. “They don’t take into account the kids that are in SAT [Scholastic Achievement Test], the kids that are ELL [English Language Learners] or multicultural languages and how they are being assessed, as well, and that is part of our test scores.” Focusing on student growth rather than their test scores could greatly benefit the
child and teacher. By using absences, surveys, and test scores as part of the evaluation, Melinda stated that the evaluation becomes an inaccurate representation of overall teacher performance.

**Observation Process within the Teacher Evaluation**

When asked about the observation process, Melinda mentioned that this is at the heart of what teachers do each day, and teachers need to be observed as much as possible for evaluators to truly understand the dynamics of the classroom. “I would be proud to see more overall weight going into observations.” The observation process started with a pre-meeting with the evaluator. During this meeting, the teachers explained the objectives of the upcoming class period and what the evaluator should expect to observe during that time period. Then, the evaluator visited the class to see the “Cadillac lesson plan” and wrote down their observations. They looked for areas of strength and areas of improvement. The evaluator and the teacher met after the observation to discuss the collected data. When Melinda spoke to her evaluator after the observation, she found the feedback helpful. She reflected on the information and used it to better her teaching practices. Even though the evaluators only visit periodically, Melinda welcomed evaluators to visit her classroom. In Melinda’s opinion, the observation process is the most valuable area of the evaluation process.

**Purpose of the Teacher Evaluation**

When asked the purpose of the evaluation system, Melinda stated that money is the primary focus. “Someone is profiting from us doing the evaluation.” She stated that districts tend to spend money on various systems or curriculum and then change after a few years, which leads to a shortage of money.
For instance, they put $8 million into Teach Scape [system to support professional development], and then they did away with that a couple of years later, and they went into [a different system] and then that was funded by another $3 million. If we are short funds in education right now, why are they wasting this money in evaluations and stuff? Who is profiting?

When Melinda talked to the union president about the evaluation, they indicated that the current evaluation system is about money and politics. In Melinda’s opinion, politicians gain financial incentives toward their campaigns when supporting the teacher evaluation system. The money is not truly benefiting the students and schools, yet provides political gain. She hoped for a new governor that focuses on strengthening education.

In Melinda’s opinion, one of the purposes of the evaluation system is to push some teachers out of the system. Many teachers have left education or retired as a result of the new evaluation system, because they did not feel appreciated nor treated like professionals. The new system seems overwhelming and unnecessary to many educators. Teachers struggle more now to get everything done for their classroom, since their focus is being pulled in many directions. “The focus is no longer on teaching kids, it is if we are effective teachers or not, and that really affects morale because most teachers put in so much time, money, effort, and everything into their profession.” Some teachers think that the district expects too much from them. “It is just so much documentation, so much uploading. You feel like you are walking on eggshells.” When veteran teachers leave the system, the district hires first year teachers or others with less experience. From Melinda’s perspective, the district then receives an opportunity to replace them with inexperienced teachers who cost less to hire, again focusing on money rather than student benefit.
Effectiveness of the Observation Instrument

When asked about the effectiveness of the observation instrument, Melinda immediately mentioned the time consuming process of uploading the needed artifacts to complete the evaluation process. Melinda spent over eight hours completing this task.

It has become an expectation now, where it is just more tedious work that we have to do to prove we are effective teachers when we should be spending that time working on how to direct instruction for our students and differentiate it and target instructions. Unfortunately, that is not what it is being used for. So, it is just so frustrating.

Even after uploading all of the artifacts, the effectiveness of the process depended on the actual evaluator. During a conversation between Melinda and an evaluator, the evaluator indicated that the marks she gave teachers during observations could depend on her mood and stress of her days. This conversation indicated that the accuracy of the observation diminishes even at the level of the evaluator. “So, it depends on what mood you are in, good or bad, or time crunch, or if you like the person when you determine the scores.” The subjectivity of the data collected during the observations led Melinda to doubt the overall accuracy of the observation instrument.

Components of the Observation Process

The relationship between the evaluator and teacher can be beneficial throughout the evaluation process. Melinda found her current principal to be very supportive during the evaluation overall and specifically during the observation portion of the evaluation. “This year, I feel a bit less heavy because we have a new administrator who is working with us, but in the past it was awful. I felt like I couldn’t teach because I felt like I was going to be dinged
for any little part of the eval.” Melinda’s current evaluation included advice to help teachers efficiently navigate the process. For example, “she says the more you write doesn’t guarantee you a higher score.” The current principal indicated that administrators can tell when they observe a “dog and pony show” and look specifically for “highly effective work done on a regular basis.” Melinda found this advice helpful. Her current principal supported her throughout the entire process.

Melinda welcomed visits to her classroom and told evaluators to come in as frequently as possible. “I really think a high weight should be on our observations because that is what we do as teachers.” Last year, due to busy evaluator schedules, evaluators did not observe classrooms as often as Melinda wished they would to best understand the dynamics that occurred within the classroom. The dynamics of a classroom include student academic performance, behavior management, accurate assessment, and many other components. There are going to be days when the dynamics of the classroom are flourishing more than other days. “You can see the times when it does not go so well, when you have tons of interruptions, or the kids are not feeling very good, or that kind of thing.” Multiple visits increase accuracy of the observation process, as well as the overall evaluation.

Melinda did not find the “Cadillac lesson plan,” which is expected as part of the evaluation system, to be a useful component. “We do a lot of lesson planning, but the way they have it weighted, I do not agree with it.” The state expects an incredibly detailed lesson plan prior to the observation, yet teachers do not have the time needed to use this type of lesson plan on a daily basis. “It is not reality to have that on a daily basis, so I feel it is kind of a waste of time for us to do but it does show everything you need to know.” Melinda stated that the rubric used for the lesson plan needed to be adjusted to better support the daily
experiences of the teachers. In addition to the lesson plan, Melinda’s evaluator asked her to provide summaries focused on different aspects of her teaching, even though this data was not required by the district, and took multiple hours to compose.

**Data Collected during the Observation Process**

In Melinda’s experience, evaluators tended to collect data during the observation to share with teachers during a post-conference. They focused on the various domains and wrote down strengths and challenges observed during the visits. During the post-conference, she received an opportunity to clarify parts of the lesson or answer any questions if needed by the evaluator. The data collected during the observation provided the teacher with information to reflect upon. For example, Melinda mentioned that her evaluator collected data about her transition times and the types of questions she asked during her lesson. Her evaluator discussed these areas with her and encouraged her to continue to reflect on these areas. Were the students given enough time to answer the proposed questions? Were the questions promoting high-level thinking? The discussion she had with her evaluator about these areas helped her reflect on her practice. This type of feedback encourages self-reflection and improvements to teaching. Melinda appreciated receiving the collected data in these areas from her evaluator. Besides the written data, Melinda was not aware of other data collected during the observation.

**Changes to the Observation Process**

In Melinda’s opinion, the observation process needs to be changed. First of all, the evaluators need to be present in classrooms more often. Melinda mentioned that evaluators need to visit classrooms more frequently to truly understand the needs of her students and the dynamics of her classroom. By increasing the number of visits, the evaluators could truly
understand the effectiveness of the strategies used in Melinda’s classroom and the impact on her students.

Secondly, Melinda also mentioned that the amount of paperwork must decrease. Teachers can’t be expected to spend hours uploading artifacts and providing reflective statements for multiple domains. She wanted to spend time focusing on her students and planning fun lessons. The time spent on paperwork took away from her focusing on more important aspects of the job that had more impact on students.

Thirdly, the information used to score teacher performance needs to be checked for accuracy. As mentioned prior, Melinda’s data was not accurate for multiple years. When evaluators use inaccurate data to judge teacher performance, the value of the evaluation diminishes in the eyes of the teacher.

Final Thoughts

Melinda expressed that the evaluation system is extremely subjective and not accurate. For example, a teacher at Melinda’s school mentored a first-year teacher, and during that school year, the mentor teacher received a minimally effective rating while the first-year teacher received exemplary scores. How does a first-year teacher receive exemplary scores? A first-year teacher does not meet the rubric standards to receive a score of five. The mentor teacher felt “demoralized” by this situation and frequently considered early retirement. The first-year teacher expressed discomfort by the situation. The system caused divisions to form among teachers. As teachers left Melinda’s current school due to the evaluation process, the overall ranking of the school went from a B to a D. At this point, Melinda was satisfied with receiving effective scores, which allowed her to avoid being
placed on an improvement plan. Her colleagues shared that the improvement plans take an abundance of time to build and orchestrate.

Melinda mentioned how the new evaluation system also negatively impacted morale. Teacher trust towards the district and each other diminished. The constant turnover at her school made it difficult to build trusting grade level teams. “We just need consistency, and we need to build trust again, and I think how we do it is through evaluations. We are okay being evaluated. That does not bother us at all. We just want to make sure, when we are evaluated, it is not skewed data.” In Melinda’s opinion, a more accurate evaluation system would promote more collaboration between teachers and the district, ultimately rebuilding trust and creating a more useful evaluation. By maintaining trusting relationships among the professionals, the students benefit.

Due to her lack of trust in the system, Melinda decided “not to care anymore.” Recently, she left a few sections blank on her evaluation, because she needed the time to complete student progress reports and lesson plans. The students and parents promptly needed the progress reports, so Melinda decided to prioritize the reports and lesson planning above the completion of a few evaluation sections. She mentioned that the district does not understand how much time is needed to complete all aspects of the evaluation process. Teachers spend multiple hours outside of the school days preparing documents. By adding the evaluation into a very full plate, teachers feel frustrated. “I feel like I really used to take time to make my teaching fun, and now it is just so calculated on teaching to the test, and I don’t agree with that.” In order to plan dynamic lessons for the students and prepare the needed artifacts for the evaluation, teachers need more time, and “there is not enough time in the day to get it all done.”
Melinda mentioned that the district office loses perspective about greater student needs.

I think a lot if the kids, when they come in, their needs are not being met, whether they are hungry, or they don’t have clothing, or their parents got in a fight that night, or someone was in jail, or domestic violence, or a lot of abuse they are having to deal with. And, here we are worried about the student test scores?

Teachers understand that their students deal with many situations outside of school that directly impact their academic performance. The evaluation currently does not take those aspects of a child’s life into consideration when scoring teachers. When students feel distress, the teacher must first support him with his emotional needs. This must occur before teaching curriculum. The evaluation must reflect the needs of the whole child and the ability for teachers to cater to those needs.
Chapter 9: Findings

This chapter focuses on the main themes that surfaced from interviews with the five study participants. The teacher grade level, teacher curricular focus, and school report card varied with each participant. Mary Gene and Catie both taught special education in elementary schools. Mary Gene’s school received an F, while Catie’s school maintained a school rating of C. Melinda taught kindergarten in a D-rated school, and Kaitlin worked with second grade students in an A-rated school. Debbie worked in a B-rated middle school and taught science and Spanish. All school rating levels (A-F) were represented.

The research focused on a central question: *What are teachers’ perspectives on the observation process within the teacher evaluation process?*

From the interview questions, data was collected and organized into themes. The participants’ responses, which focused on their perspectives of the observation component of the teacher evaluation process, provided four distinct themes. In reflection of a phenomenological approach, the themes illustrated the teachers’ complete, lived experiences with the teacher evaluation system.

There were four themes that surfaced from the data analysis: a) Theme 1: The Relationship between the Teacher and Evaluator, b) Theme 2: Teachers Respond with Frustration Due to a Lack of Purpose, c) Theme 3: Teachers Question the Accuracy of Observation Data, and d) Theme 4: Evaluators Need to Know Their Students.

**Theme 1: The Relationship between the Teacher and Evaluator**

During the observation portion of the evaluation system, the participants expressed the need to receive guidance from evaluators about how to improve their teaching and better support students. They also mentioned how they appreciate constructive feedback that
promotes growth, during the observation process. They work in the public school system because they care about their students and want to support them to the best of their ability. Receiving feedback and learning new skills and strategies better serves the students. During the observation portion of the evaluation, teachers’ relationships with evaluators impact the provided feedback and the written evaluation. All of the participants commented on the impact their relationships with evaluators had on the content, and the results of the actual teacher evaluation.

Debbie mentioned that if a teacher is in the “cool club,” then the teacher received a better evaluation. A teacher who did not have a positive relationship with the evaluator or was not in the “cool club” might have to work harder and “push extra hard to earn the proper evaluation that you need.” Debbie’s last evaluator was a close friend of hers, which led the evaluator to tell Debbie that she did not need to observe her classroom because she trusted that Debbie did a good job. Debbie received all 4s and 5s on her evaluation. Debbie felt disappointed that her friend did not come in her classroom to observe and evaluate her performance. “I know this isn’t right, because I know she did go to other classes, so how did those teachers fare? So, you know there is a little bit of bias in terms of that.” Debbie did not approach the evaluator to express her thoughts because her scores were so high. “If you get an administrator that you really connect well with, you might get better scores versus an administrator that you don’t connect well with.” Debbie felt that her connection with the evaluator led him to give her high scores without even coming to watch her teach.

In Melinda’s experience, getting along well with her evaluator positively influenced her evaluation. She had an evaluator who allowed her personal relationships with teachers to impact the actual scores on the written evaluations.
With our previous principal, if you are on her good side, you got a good score, and if there is anything you do not agree with them on, then it is a way for them to ding you where you feel like your license is at jeopardy.

Melinda felt that all of her hard work might not have been reflected in her evaluation if her relationship with the evaluator was not strong. Whether the evaluator “likes the person or not can determine the scores, and that is very frustrating.” In Melinda’s opinion, the teacher’s relationship with the evaluator should not directly impact the authenticity of the evaluation scores. This reduces the overall consistency of the scores provided to teachers throughout the evaluation process.

Mary Gene questioned the evaluation system due to the ways different evaluators let their personal relationships with teachers alter the evaluation scores. She mentioned that administrators used the entire evaluation system, not only the observation portion, to keep the teachers that they liked best at their school.

Administrators are stressed out, too. You are going to keep those you like, but you don’t who you don’t, and some people are going to get extra points because you like them and some people aren’t. It is just not an appropriate system.

Because evaluators treated teachers differently depending on their relationships, Mary Gene did not see the value in the data collected from the evaluators. She believed that the system was dysfunctional and did not truly represent teacher performance. Mary Gene also thought that administrators assigned their favorite teachers’ classes with easier students each year. As a result, the students were easier to teach, which also led those teachers to receive stronger scores during observations. “The administrators have their favorites, so they will give them all the kids that don’t have any behavior issues and then others, because they don’t
like them, get all the problem kids.” When administrators group students by behavior and provide specific students to teachers whom with they had positive relationships, it alters the observation portion of the evaluation system, as well as other areas.

Catie frequently received exemplar marks on her teacher evaluations. However, when a new principal arrived, and Catie asked questions about parts of the evaluation, the principal got “very defensive.” Catie asked questions because the evaluator had limited experience with special education, and Catie hoped to collaborate on the evaluation to make sure the evaluator understood the daily needs of her students. Catie and her new evaluator did not get along, which negatively impacted Catie’s observation scores and overall evaluation. “She and I have had issues.” Catie no longer received exemplar marks even though she taught the same group of students and continued to use similar strategies. As a special education teacher, Catie felt that the evaluator did not understand some of her day-to-day challenges, which caused friction in their relationship. “I argue with my current principal all the time.” Catie tried to explain her daily challenges to help the evaluator better understand her students and classroom environment, but by being vocal, her relationship with her evaluator soured. Catie did not feel supported and thought the data collected during observations did not accurately represent her performance. From Catie’s perspective, her negative interactions and poor relationship with her evaluator led her to receive lower marks on all areas of the evaluation.

Kaitlin developed strong relationships with her evaluators. “I think that I am lucky, and I have an administrator that I personally get along with really well, and she gives me good feedback.” Since Kaitlin had been evaluated by the same person for a few years and
had a strong relationship with her, she thought that the feedback during observations had not changed.

However, going back to the fact that I have had her so long now, I do think that she kind of knows what to expect and maybe only half listens because she knows I am strong and kind of does whatever. I was actually hopeful that I would get the new principal this year because I thought this might be fun to get some new feedback. Though Kaitlin’s observation feedback had been positive and helpful through the years, she felt that now, due to the positive relationship with the evaluator, she continued to receive positive marks at the expense of receiving new feedback. Because her evaluator developed specific judgments about Kaitlin’s teaching, she provided similar feedback each time. In Kaitlin’s perspective, her positive relationship with her evaluator has kept her from receiving new and potentially helpful feedback. She hoped to work with a new evaluator in order to receive new feedback and to continue to grow professionally.

In summary, when all five of the teachers were asked about their perspectives of the observation process within the teacher evaluation system, they each mentioned an impact on their evaluation scores and feedback that resulted from their relationship with the evaluators.

Theme 2: Teachers Respond with Frustration Due to the Lack of Purpose

Overall, participants were angry and frustrated about the unclear purpose and lack of effectiveness of the observation portion of the evaluation system. When the purpose is unclear, people get frustrated. The participants mentioned the importance of observations for professional growth, but also indicated that there were limited or no resources for growth in the current evaluation system. Some participants mentioned that the new system has led effective teachers to retire early or leave the school system. Also, school climates and
All of the participants expressed anger and frustration when asked about the observation process within the teacher evaluation system, as well as concerns about the larger impact on teachers and the purpose of the system overall.

Mary Gene stated that the evaluation system “de-professionalized” teachers and placed them in impossible situations. The observation portion of the evaluation system should focus on mentoring, improving teacher performance, and professional growth. However, the current system does not focus on growth or improvement. Instead, Mary Gene stated that this system does not support teachers at all. “The whole thing is ‘how are the kids going to perform on [state standardized testing] days,’ and I got a bad evaluation so what do I do now?” Mary Gene mentioned that teachers are highly trained to work with students and are not being trusted to do so. She mentioned that she feels trapped within her profession and the negative evaluation system. “I went to school for nothing, so there is nowhere I can go.” The evaluation system does not focus on helping students and teachers, and in some ways does the complete opposite. “You can go to an A school and do well for yourself or you can go to a school that needs you and screw yourself.” In Mary Gene’s opinion, the evaluation system encourages teachers to work at schools with high report card scores. The higher performing students naturally are easier to teach, and teachers’ observation data will be higher, regardless of the strategies used in the classroom.

Melinda stated that she “is so upset and frustrated” and “not appreciated” when reflecting on her experience with the observation process within the teacher evaluation system.
system. Teachers are not treated professionally and, as a result, “we have lost a lot of good
teachers over the evaluation system they have in place.” According to Melinda, the current
teacher observation system increased teacher workloads. The district requires teachers to
turn in additional paperwork. Teachers are currently leaving the system due to the evaluation
system and stress involved. The current evaluation “deters new teachers from staying in the
system” and makes veteran teachers “feel offended by it and want to leave the profession
because it is so much.” The high turnover is not helping increase student achievement.
Melinda mentioned that overall teacher morale has been negatively impacted by the
evaluation system because teachers spend so much time, effort, and money on their
profession and are not treated professionally. Melinda commented that the current system is
“divisive.” “It causes division within the staff as opposed to team building.” Melinda
mentioned that she does not think that the time and effort spent on the observation portion of
the evaluation system is making her a better teacher, because there is a lack of purpose and
the current feedback is not helpful to her professional growth. It is taking time away from
her students and community building in schools, which should be the primary focus. “Why
am I doing so much? At this point, it is too much.” Melinda stated that the observations and
the entire evaluation system are exhausting. As a result, it is difficult for her to care any
longer about her observation data and scores.

Kaitlin also commented on negative reactions in regard to the observation process
within the evaluation system from teachers at her school. “A lot of people I know are hostile
about the whole thing because it looks so different, and I feel they don’t like change.”
Kaitlin mentioned that she hopes the purpose of the evaluation is to help teachers improve
but is concerned that effective teachers are leaving the system due to the changes, which ultimately negatively impact the student experience.

What it is doing is driving out good people, and people are retiring early and that is probably one of my biggest concerns about it. I feel like all of these unintended consequences that we won’t see until years down the road where we are spending money to replace these educators who are good and now we have less effective people there.

Kaitlin stated that the current evaluation system is pushing good teachers to leave, when the purpose should be to help teachers improve. She understood that ineffective teachers might leave the system with a new evaluation system. However, the teachers that Kaitlin observed leaving are effective and will be missed. “It is hard not imaging that they aren’t doing it to get rid of ineffective educators, but that is not essentially happening anyway.” This type of evaluation is all Kaitlin has ever known, so she has not experienced the large change teachers constantly discuss. However, even her initial neutral view on the system had been impacted by listening to other teachers’ negative reactions and seeing so many people leave the profession. “I am tainted though by listening to the naysayers and the people that complain about it and can’t stand it.” Kaitlin observed effective teachers leave, and she questioned the purpose and effectiveness of the observations and the entire evaluation system.

Catie, whose scores have changed from *highly effective* to *minimally effective*, expressed her negative feeling toward the evaluation. Catie believed that evaluations are necessary to improve education and promote growth. However, in her experience, this system is doing the opposite. “It is incredibly punitive. Every year, I feel more beaten down by this.” Teachers embrace their profession because of the “pure intrinsic” value, not
because of monetary value. Catie cried multiple times when she talked to her evaluator about the data compiled from the observations and overall scores. She put her heart and soul into teaching her students, and she believes that the system is not about improvement, but instead makes teachers feel inadequate. “You are being told something opposed to you being encouraged to change or to improve, to grow.” As a result, Catie shut down completely.

When her evaluator asked her to complete a lesson plan before the observations and other tasks for the evaluation, Catie asked about the purpose of the task. “I asked why.” The evaluator mentioned that she would have to reduce points if Catie did not complete all of the tasks. Catie did not see the purpose of spending hours completing tasks that wouldn’t make much of a difference on her evaluation or for her students’ experiences in the classroom.

I will never stop doing the effort for my kids. I will never stop doing what I know is right for them, but I will stop putting on a performance or a show for somebody else who doesn’t know my kids, who doesn’t know why I am doing what I am doing.

Because Catie was not the only teacher at her school that felt angry, the overall implementation of the evaluation negatively impacted her current school climate.

Debbie also expressed frustration with the evaluation system. She thinks that the primary purpose of improving teaching is not effectively being achieved. “I am not quite sure what the end goal is.” The goal of the evaluation system seemed unclear.

If the goal is to make classrooms better, teachers better, student learning better, then I would say it isn’t effective at all because there is no follow through with these evaluations. We get our score and that is it. There is nothing that says there are these opportunities out here.
When teachers receive scores, Debbie stated that teachers then need to receive feedback and ideas about how to improve their practices. However, the evaluators did not spend time analyzing the feedback and helping teachers grow. “The evaluation is just there. It is just another hoop to jump through for most people.” It appeared like an additional task, which decreased the overall effectiveness. When the purpose is unclear, teachers feel frustrated and unacknowledged for their efforts. “I have seen teachers that do feel unappreciated and have left because of that, and I don’t blame them.” Teachers left the profession due to the evaluation system and its lack of purpose. Debbie understood that observation and evaluation are necessary, yet did not think the current evaluation clearly encouraged teacher growth and improvement. If the purpose is more clearly stated and implemented, then overall effectiveness would increase and low performing teachers would leave the field due to the process.

Debbie and Kaitlin, who are at the two schools with the highest report card scores, both expressed frustration with the unclear purpose of the evaluation and the lack of effectiveness towards achieving that purpose. They both discussed their concerns about losing effective teachers and how the evaluation negatively impacted overall morale. However, as they both consistently received high marks, they both were less concerned for themselves, yet expressed how the overall system had a negative impact on their feelings and made them doubt the effectiveness of the evaluation. They expressed empathy toward the other teachers’ situations and understood why teachers left.

In contrast, the other three teachers who worked at struggling schools had a different experience. They are the teachers who are considering leaving the system. Mary Gene felt trapped because she does not have options in other industries. Catie stopped engaging in the
evaluation system because she did not see the value and willingly took the consequences. Melinda felt overwhelmed by being treated unprofessionally. As all five teachers questioned the purpose and effectiveness of the evaluation, Mary Gene, Catie, and Melinda were more directly impacted by the changes. They all three received high marks at the beginning of the new process and, with changes in evaluators, received lower marks in recent years even though they spent hours completing the required steps. Each stated that they do not think the current system is helping them improve, though it takes hours away from their students. These three teachers questioned whether or not they would stay in the education system at all. All participants discussed negative feelings about the observation process and the overall evaluation system however those at high-scoring schools appeared less affected by all the systematic changes compared to teachers at lower performing schools.

In summary, participants were angry and frustrated about the unclear purpose and lack of effectiveness of the observation portion of the evaluation system.

**Theme 3: Teachers Question the Accuracy of Observation Data**

When asked about the observation portion of the evaluation system, the participants expressed the need to increase the accuracy of the collected data. They mentioned how they understood the need for observations and would embrace the opportunity if they felt more confident in the accuracy of the information. They wanted to improve their practices and realized that classroom visits could help in that process. However, they questioned the accuracy of the data collected during the classroom visits, which decreased the effectiveness of the entire evaluation system. The participants thought that increasing the amount of time spent in the classrooms, utilizing multiple evaluators, and engaging in conferences to discuss the collected data would improve the accuracy of the classroom observation visits.
Mary Gene did not think that classroom observations showed an accurate depiction of daily classroom events. Evaluators observed teachers once or twice a year and, at her school, teachers often sent certain students out of the classroom on observation days. “We send the out the bad kids.” When the challenging students were out of the classroom, the teacher could more effectively teach the students, which did not show what truly occurred on other days. Mary Gene also mentioned that when evaluators entered the room, their presence alone immediately altered the classroom environment. “You put one thing in the classroom that is different and the whole thing changes, so you don’t really see it.” Mary Gene understood that evaluators needed to assess the classroom environment as well as the strategies used by the teacher, but she thought placing a camera in the classroom might be a better option. With a camera, Mary Gene believed that evaluators could more accurately observe the interactions between the students and teacher, and the environment would remain more consistent.

Mary Gene also believed that for an observation day, most people could create a lesson plan that would make them appear competent, and that they could remain on their best behavior with the students. “You are not really seeing how a teacher really talks to their kids. You are just kind of seeing a ‘dog and pony’ show.” After observing a classroom for only one day, how is an evaluator really going to have enough information to accurately access the effectiveness of the teacher? Teachers “shift whatever they are doing and pretend this is what you were doing all along” when an evaluator observes their classroom. In order to increase the accuracy of the observations, Mary Gene thought that evaluators should visit classrooms more often and unexpectedly. They will get a more authentic picture about what is actually occurring in the classroom.
Maybe multiple observations, maybe not planned observations, so when you come in and see what someone in doing, you get a real picture. And I know teachers would not necessarily like that and you could be having a bad day, but it could be enough times that you know that wasn’t it.

Multiple, unexpected visits may allow students to get better acquainted with the idea of having evaluators in the classroom. More visits also could help evaluators better understand what truly occurs in the classroom on a daily basis.

Melinda commented that, because evaluators are human, the assessment during classroom visits sometimes depended on the current mood of the evaluator. During a conversation Melinda had with her evaluator, the evaluator mentioned “how it [the observation data] can be skewed on her [the evaluator] level.” This evaluator indicated that scores were probably impacted when “things went south in her life, where she had a lot of things going on and things got busy at school and stressful, so the next couple of days she put good on everything because she didn’t want to deal with that.” Melinda felt frustrated when her evaluator admitted scoring some evaluations inaccurately because of her own life stresses and situations. These life situations affected her mood, which impacted the scores given to some of the teachers at the school. “Once again, it is very subjective.” Because the classroom visits are scheduled and only occur periodically, Melinda stated that they are “a ‘dog and pony’ show” and not a true representation her day-to-day challenges, which reduces the overall accuracy of the observation and potential results.

Melinda invited multiple evaluators to visit her class on several occasions for a prolonged amount of time, to truly learn about the dynamics of her classroom. She was confident in her practices and hoped to demonstrate them to various evaluators. If evaluators
visited more frequently, the collected data would be more meaningful. By engaging in more unexpected visits, the evaluator would receive the opportunity to “see the times when it does not go so well, when we have tons of interruptions, or the kids are not feeling very good, or that kind of thing, and seeing how the teacher handles it and how they are able to move forward.” Teachers deal with multiple obstacles every day and increasing visits provides evaluators with more information and a more accurate understanding of the classroom dynamics. Melinda understood the importance of evaluation and wanted to receive the data, yet expected a certain level of consistency. “We are okay being evaluated. That does not bother us at all. We just want to make sure when we are evaluated, it is not skewed data.” When teachers receive inaccurate data, trust between teachers and administrators diminishes.

Kaitlin commented on the need for evaluators to spend more time in the classroom. Evaluators cannot authentically understand the effectiveness of a teacher without observing the teacher handling multiple situations with students. “I don’t think anybody can see a person’s whole gamut of teaching in just an hour observation.” Kaitlin understood the need for observations, however, she thought that the collected data would be more accurate if the evaluator visited more often and stayed for longer periods of time.

I feel like that could happen more often, and that might give a little better indication of what is going on in the classroom, their true dynamics. Maybe stay a little bit longer. You know, anyone can pull anything off for five minutes, but if you stay for 15 and you might see a whole different process.

She also believed that, sometimes, when using checklists during the classroom observation, the evaluators inaccurately made assumptions. It is important that the evaluator and teacher have time to collaborate to increase the accuracy of the collected data. For example, Kaitlin
had rules hanging in the classroom and the evaluator assumed the students did not have a role in making the rules. Kaitlin got reduced scores for not incorporating student feedback on the classroom rules even though the students made all of the rules. The evaluator made an inaccurate assumption, Kaitlin asked about it, and it got corrected. Kaitlin truly enjoyed sharing her classrooms experiences with an evaluator, yet she preferred that the evaluator be less focused on completing a checklist and would participate more with the students. “I wish it wasn’t all them sitting there checking boxes. It would be nice if they could just sit there and enjoy and take it all in for a minute.” If the evaluator took more time being present in the classroom or discussed the data before or after the evaluation, they would understand the true dynamics of the classroom and the class culture.

Due to her negative experiences, Catie questioned the accuracy of the observation process within the evaluation system. Even though she worked with the same students using the same strategies, Catie’s scores changed from exemplary to minimally effective. As a special education teacher, evaluators needed to understand the challenges faced in Catie’s classroom and how they differ from general education environments. Catie’s prior evaluator understood the nature of a special education classroom, yet her current evaluator has limited knowledge. Because her current evaluator “doesn’t know what to look for” during an observation, Catie’s scores declined. Her prior evaluator visited her classroom often and worked with Catie for three years. She knew the effectiveness of Catie’s practices. Her current evaluator rarely visited the classroom. “For my observation, she came into my classroom 30 minutes late and then marked me down on things she didn’t see because she wasn’t there.” Catie wanted to improve and grow, yet her current evaluator’s lack of experience decreased the accuracy of the collected data. “I have been raked over the coals.
This supposedly objective system is completely subjective.” Catie thought that her scores fell because her current evaluator did not understand the complexities of Catie’s classroom, which led Catie to think that the process was completely subjective. “I think principals are rushed. They are trying to do one-hour evaluations on everybody. They come in, and they have their paper, and they are looking for things.” Evaluators move quickly through classrooms, yet only absorb a limited amount of information to utilize when producing teachers’ scores. When Catie felt she received inaccurate scores, she presented her reasoning to the evaluator, but the evaluator never changed her scores as a result of those conversations. Sometimes, Catie received lower scores in an area because the evaluator was not present to assess a specific skill. Catie believed in the importance of observations, but she did not think that the current system provided accurate information. In order to increase accuracy, in Catie’s opinion, evaluators need to observe more often and pop in to classrooms unexpectedly. “I tell her to come in whenever you want.” Catie wants to receive accurate data to help her improve her practices, yet at this point, the classroom observations do not accurately depict the effectiveness of the teacher because they are planned and limited. “I refuse to put on a ‘dog and pony’ show.” Catie did not think that the observation process helped her improve her practices.

Debbie questioned the accuracy of the observation system because only one evaluator visited each year and collected data through her specific lens. Also, evaluators tended to visit for only part of a classroom period, missing either the beginning or end, which led them to collect limited data. Evaluators also did not often meet before or after the evaluations to discuss the data and check for accuracy.
So, since it is not committee style, it is one person saying I see this, I see that, for 10 and 15 minutes. I think this past year, I actually had an administrator sit in my class for 30 minutes, which a whole class is over 50 minutes, so really they have never been there for a whole class period, and once, maybe twice, that I got evaluated. I don’t feel like it is an accurate description of what I am doing.

Due to the time constraints of the evaluators, teachers often do not have the opportunity to meet to discuss the collected observation data, which leads to inaccuracies and misunderstandings. During her observation, Debbie commented that “there wasn’t an official anything. We didn’t talk about it before hand. We didn’t talk about it after hand. I just got some scores online.” Meeting with evaluators before and after the evaluation would increase the accuracy of the data and hopefully lead to reflective conversations.

In Debbie’s opinion, a primary way to increase accuracy is engaging more evaluators in the process. “I don’t think you can get an accurate evaluation from one person, two days out of the year.” Teachers could benefit from working with more than one evaluator. By compiling multiple evaluations, the data becomes more accurate. The patterns found in various evaluations could lead to rich conversations between the evaluators and teachers. “You should have someone who is in your department, so they can understand. You should have someone who is maybe a little senior over you in terms of your teaching experience, maybe an administrator as well, and maybe someone else that you are working with on a day-to-day basis like a previous mentor or someone else that really knows you.” Debbie stated that working with more evaluators provided more accurate, thorough information to increase teacher reflection and growth. Also, working with evaluators who are familiar with your teaching style and strategies could also lead to more accurate data and scores.
Debbie mentioned that increasing the number of visits to the classrooms also provided more accurate information in the observation portion of the evaluation system. I think also having more days that you are evaluated, maybe instead of just a single day having an entire week or an entire unit or something like that, so they can see that within one unit you are really hitting all of these different points. It is really hard to gauge in one day what you are doing. Making it a bit longer. I don’t mind being evaluated. I like it actually. I just wish for a more accurate representation of what is happening.

When evaluators visit classes frequently, they better understand the relationship between the teacher and the students within the classroom, observe various situations and the way the teacher reacts, and grasp the class culture and morale. These classroom qualities are hard to grasp with limited exposure. Evaluators currently collect data and score teachers in a variety of areas after spending limited time in the classroom. Often, teachers receive scores in areas that were not observed during a limited evaluation timeframe.

Because Debbie did not think that the observation part of the evaluation system accurately represented her ability, she mentioned that, in her opinion, the objective of the evaluation system must be to make sure teachers fulfill the basic needs of the students. The system does not focus on teacher growth and reflection. “I can’t accurately say that it is going to give a good representation. It is just a way to show that we are doing things in the classroom, essentially. That is what it feels like because I don’t feel like I am being accurately represented.” Without accurate information, the evaluator does not know what growth needs to happen in the classroom, and the teacher can’t use the data for reflection. “If the goal to make classrooms better, teachers better, student learning better, then I would say
TEACHERS’ PERSPECTIVES OF THE OBSERVATION PROCESS

it isn’t effective at all because there is no follow through with these evaluations.” Debbie assumed the goal of the observations was to help teachers grow, but surmised that was not happening and can’t occur until the accuracy of the evaluation increases.

In summary, when all five teachers who participated in the study were asked about their perspectives of the observation process within the teacher evaluation system, they mentioned the need for evaluators to spend more time in the classrooms to better understand the aspects of the classroom environment. In order to score teachers on multiple areas, evaluators need to spend more time in the classrooms to collect authentic data, which will ultimately increase the overall accuracy of the evaluation system.

Catie and Mary Gene mentioned the importance of unexpected classroom visits. Melinda and Debbie felt that multiple evaluators could help increase the accuracy of the data. Kaitlin and Debbie indicated that conferences before and after the evaluations could improve accuracy.

**Theme 4: Evaluators Need to Know Their Students**

The participants expressed the need for evaluators to know the specific needs of the students in their classrooms before engaging in the observation process. Some external factors in students’ lives impact the way they act and engage in classroom activities. If the evaluator understands the background of the students, they can better understand the strategies being used in the classroom. Sadly, some of the students are homeless, hungry, in homes experiencing domestic violence, or dealing with chronic illness. Teachers use different strategies with different students, depending on their specific needs. Evaluators must understand the challenges faced by the teacher and students in the classroom and take them into account. If these variables are not considered, teachers may ultimately opt to
migrate to more affluent areas of the community or in grade levels with less measures of accountability (e.g. kindergarten does not take state standardized tests). Subgroups of students in the public schools need different support from teachers, so evaluators need to be made aware of those population challenges in order to understand how it could affect the teaching environment and strategies, as well as the data collected during the observation portion of the evaluation system.

All the participants commented on the need for evaluators to understand the needs of the students in the classroom. Different grade levels and subgroups have diverse needs, which impact the ways teachers approach their communities of students. In order for evaluators to collect accurate data and understand the strategies implemented in the classroom, they must first know the dynamics of the classroom and students involved. Because the evaluation system focuses only on all students achieving grade level, teachers are not scored or credited on each child’s growth.

Mary Gene, Melinda, and Catie spent a large portion of their interviews talking about the evaluators needing to know the students and considering this information when deciding on scores. Mary Gene stated that many of her students have not learned basic skills that are necessary for productive learning in the classroom. These needs must be fulfilled first in order for the students to be productive in the classroom. She used class time to help students develop these essential tools.

My kids are like little beasts; I mean they have never learned how to sit, how to say ‘thank you,’ how to do anything, so my whole thing has been behavior. Of course the academics are there, but it is keeping one kid from killing another kid and learning how to sit and listen, and it is coming along and people notice. People say ‘these kids
are night and day from where they were’, but you are at a place where you are trying
to teach a kid to sit, so how are you going to do this academic, fabulous thing when
your kids don’t know how to do that? They don’t have the home support.

Many of her students received no educational resources at home and, as a result, Mary Gene
had to teach them additional skills. Because the support at home was limited, Mary Gene
taught them skills that students in other areas of the city would have already mastered upon
arriving to school.

Mary Gene stated that her performance should be based on student growth, not
achieving grade level performance. “They don’t take into consideration your kids and their
levels, and it’s just like you are supposed to be doing this, and this is what we are looking
for.” Her students learned so much about social skills and how to be part of a community
when they arrived to her classroom. If the evaluators knew her students, they would
understand the amazing growth made by the students each year.

You need to look at where a kid is and where they go. You don’t need to look at if
they are at grade level. We don’t look at growth; we look at grade level. I have kids
who have made two to three years growth, but they aren’t on grade level…kills me,
because it doesn’t show. It doesn’t show because they don’t know where they come
from so they don’t know where they are. Nobody knows where the kids are except
their teachers.

When evaluators visited the classrooms, they did not understand the specific needs of the
students and judged the classroom environment with limited knowledge and understanding.

Mary Gene shared her specific challenges with a student this past school year.
We have a little boy who had more than 200 [physical outbursts during a school year] attacks [towards students and teachers], and he has fingernails. I had to cut his fingernails in class because he scratched up so many kids in class. He is on the spectrum, but he is doing great now.

Due to the challenges they face, achieving grade level performance is extremely challenging and, in some cases, unachievable for certain students. Working with students that have additional challenges can negatively impact the data collected during the observations. Evaluators do not always understand the growth achieved by students nor understand specific strategies used by teachers to support those students. When students’ basic needs are not met at home and parents provide limited support, the teacher picks up the additional challenges to help students grow.

You are at [a high school in the Southwest] and the kids [at the high school], not all of them clearly, have had family support, and they are well fed, and they sleep. My kids don’t have enough food, they don’t have a home, they don’t have parents that know how to parent, they have parents who are working all of the time…There has to be some understanding in the evaluation process of the students, not just the teachers but the students, and the parents, and the home and poverty and all this stuff doesn’t come into play. It is about the teachers either failing or not. You can go to an A school and do well for yourself or you can go to a school that needs you and screw yourself.

In Mary Gene’s opinion, the observation portion of the evaluation system did not take into consideration the specific students supported in her classroom. If teachers were assessed on student growth and evaluators knew the challenges of the students in the classroom, the
observation portion of the evaluation system would more accurately assess teacher performance and encourage growth.

Melinda, similar to Mary Gene, commented that the observation part of the evaluation system should make adjustments to teacher scores depending on which subgroups of students the teachers support. Students that live in various areas of the city face diverse challenges and have different needs. Poverty affects great swaths of the population in certain areas and one in four children are going hungry, on top of other learning or developmental disabilities. This information should factor into evaluations and scores. The evaluators need to know the students in order to accurately collect classroom data and score teachers.

They don’t take into account the kids that are in SAT [Scholastic Achievement Test], the kids that are in ELL [English Language Learner] or multicultural languages, and they are being assessed, as well, and that is part of our test scores, and they don’t see where these kids came from and where they exited at the end of the year. They should look at student growth rather than student test scores. That is what I think because, when we see that a lot of them have not been in early childhood programs or preschool, you know that kind of thing so when we get them. They have absolutely zero knowledge, but then you have a few families who have worked with them or they have been in childcare, you can see their growth.

By focusing on growth instead of being on grade level, teacher scores will not be as negatively affected by the fact that they work with specific groups of students. If evaluators understand each student’s growth, then they can collect more accurate data during observations and better support the teachers.
Making sure that the basic needs of children are met is Melinda’s primary concern. The range of challenges faced by children is huge in this southwestern city. Some are in violent homes, while others do not know where their next meal will come from. When basic needs are not being met, it is challenging for a student to focus during class or while taking tests. Maslow’s Hierarchy of Needs theory depicts this issue starkly. If there is a failure to meet basic needs, such as food and safety, people have difficulty moving beyond that to learn or grow.

I think a lot of kids, when they come in, their needs are not being met, whether they are hungry or they don’t have clothing, or their parents got in a fight that night, or someone was in jail, or domestic violence, or a lot of abuse they are having to deal with, and here I am worried about their test scores….So, I say if their test scores don’t get there, so what. I will take a hit for it, but their needs need to be met before I can even teach them.

When many students have limited resources at home, Melinda stated that it creates additional challenges in the classroom that will most likely be observed by the evaluators. The evaluators need to understand the students’ situations and incorporate this knowledge into the data and teacher scores. Otherwise, teachers are going to be compelled to move to other areas of the city for employment, where students are being educated by parents and their basic needs are met. Due to the additional resources of those students, it is more likely that they will focus better in class and achieve grade-level performance.

Similar to both Mary Gene and Melinda, Catie thought that evaluators needed to know the students being taught in each classroom and their specific challenges. Evaluators should be looking for student growth when observing in the classroom. Catie worked with
students who might never achieve grade level status due to the nature of their specific learning challenges. However, the evaluators did not always understand the specific challenges these particular students faced, or the amazing growth that occurred. The evaluators focused on their checklists and whether or not specific criteria were fulfilled during the observation.

I have a kid who I joke is the only kid who could fall off the floor, because he is so ADHD along with autism and now he is sitting in general lab for language arts. Is he doing well on [the state standardized] test? No, but he has shown progress, but they don’t care about that. All they care about is that he is still under grade level. Because evaluators were unfamiliar with the students, they oftentimes collect incorrect data and, due to their lack of knowledge, scored teachers inaccurately.

I have a family, bless their hearts, they were told their child would never read, and he is reading. At the last IEP, the dad cried, and normally he is so stoic. He said, ‘They told me my son would never read, and you have him reading, and that is amazing,’ but I am minimally effective. Because Catie’s current evaluator did not understand the challenges faced in her classroom each day, her observation scores changed dramatically when the new evaluator arrived.

Debbie and Kaitlin, who teach at A-ranked and B-ranked schools, commented less on the importance of the evaluator understanding. They both recognized that their students have different needs than students at C-, D-, or F-rated schools. However, Debbie and Kaitlin both stated that they did not think the evaluators truly understood the needs of their students and their learning either. Debbie indicated that the evaluator “came in for 20 minutes, every now and then she’ll pop her head in, but nothing that I think really gauges the students’
learning.” Debbie thought that it could be valuable for the evaluator to better understand the needs of the students and their student learning and did not think that was achieved during short, sporadic visits. Debbie stated, “You are not getting an accurate representation of what lies in class.” Kaitlin thought that the evaluators needed to get to know her students by spending time with them.

I think that before learning can occur, you have to build a relationship with them, and I feel like I am really good at that, and for me, I would really like for her [evaluator] to spend more time doing that.

This would allow the evaluator to collect better data and more accurate scores. In Kaitlin’s classroom, students performed at high levels, and she did not think that the evaluators always understood the student’s capabilities. Oftentimes, the assessments also capped the scores that demonstrate the capabilities of the students.

We were not allowed to go greater than 38 in second grade, so we couldn’t show, especially in our area where we live and our kids come in strong in reading, you are not doing a lot of growth when they put a cap on where you can test up to….And, now we are using the I-Ready reading test and that does allow for us to show a little bit more growth.

Kaitlin mentioned that, even though her students perform well, it was important that the evaluator understood their needs in the classroom as well as during standardized testing. Evaluators needed to prioritize getting to know the students in order to accurately assess the classroom environment.

In summary, all five participants mentioned the importance of knowing the students in the classrooms before being able to accurately observe the classrooms. Different
subgroups of students have various needs, and the evaluators need to obtain that knowledge (McGreal, 1983; Hunter 1982). Mary Gene, Melinda, and Catie directly saw negative impacts in the observation portion of their evaluation when an evaluator did not understand the specific needs of their students. Evaluators misunderstand specific techniques used in the classroom and, in their opinion, inaccurately scored the teachers. Debbie and Kaitlin also mentioned that the evaluators need to know the students, however, have seen less direct impact on their specific evaluations, partially because they work in A and B graded schools. By understanding the needs of the students, the evaluators will better understand the dynamics of the classroom and the strategies used by the teachers.
Chapter 10: Discussion

The central research question of this study is: *What are teachers’ perspectives on the current observation portion of the evaluation process?* The participants expressed varying perspectives about the current observation portion of the evaluation process, however, their perspectives aligned in many key areas. From the participants’ perspectives, there is a need to be observed by evaluators, yet a few key changes can increase the overall effectiveness of this observation system.

**Participants Perceptions of the Impact of Their Relationships with Evaluators**

All the participants mentioned that they thought their relationships with the evaluators impacted the observation portion of the evaluation system. Whether the relationship was positive or negative, it impacted the feedback provided, the written evaluation, and final scores.

Participants who had positive relationships with evaluators reported that they received higher scores than some of their peers, though these participants did not completely understand the content behind their scores. Because of the positive relationships, the evaluators trusted that those teachers were working with students in an effective manner. In one case, the evaluator did not observe the teacher at all, yet provided feedback and scores in the 4s and 5s. These participants hoped for more valuable, content-specific feedback that could help them grow and improve their practices. They also expected evaluators to use the evaluation tools consistently with every teacher, no matter the relationship. As indicated by Pitler (2007), the tools must be used consistently for any value to surface from the data.

Participants with challenging relationships with the evaluators expressed how these relationships negatively impacted their feedback and scores. When these participants tried to
talk to the evaluators about the observations, the evaluators did not actively engage. When evaluators and teachers engage in conferences, teachers reflect on their practices and continue to grow by looking at the collected data (Goldhammer, 1969; Hunter, 1982). These participants stated that they hoped to confer with the evaluators in order to improve the accuracy of the data. They also said that their hard work was not reflected in their evaluation, which reduced the authenticity of their scores.

When evaluated by school leaders who had positive relationships with them, two of the participants received dramatically different scores. When personal relationships impacted feedback and scores, the participants stated that it reduced the value of the data and created inconsistency within the entire evaluation process. When data is inaccurate, the process does not meet the needs of the teacher nor help them improve their practices (Weisberg et al., 2009)

In summary, the participants stated that their relationships with evaluators impacted the accuracy of the scores provided during the observation portion of the evaluation system. Trust issues also tended to impact the outcomes of the evaluation (Tschannen-Moran & Gareis, 2015).

**Participants Express Frustration and Anger When Talking About Their Observations**

When asked about their perceptions of the observation portion of the evaluation system, all five participants expressed frustration about the observation process. They all mentioned that the objectives of the evaluation seemed unclear, which increased overall frustration at their schools. This frustration negatively impacted school climates. Setting a clear objective during the evaluation system helps teachers grow, stay focused, and take ownership in the system (Hunter, 1982; Marshall, 2014a; Tyler, 1950). The participants
understood and embraced the need to be observed. However, they thought that the current evaluation’s purpose seemed unclear and ineffective.

The participants reported that the new evaluation system led many effective teachers to leave the education system. Because the new system increased teacher workloads and offered unclear goals, many teachers felt frustrated and left. They did not think that they were being treated professionally.

In summary, the emotional reaction from all of the participants stemmed from the unclear purpose of the current evaluation process, the lack of effectiveness, and the increased workload experienced by the participants.

**Participants Question the Accuracy of Classroom Visits**

When asked about their perspectives of the observation process within the evaluation system, all of the participants mentioned their concerns about the accuracy of the data collected during observations. In order for the teacher evaluation system to help teachers grow professionally, the data collected must be accurate (Darling-Hammond & McLaughlin, 1995). The participants believed that the accuracy needs to improve. Methods of increasing the accuracy, as suggested by the participants, included increasing the amount of time evaluators spend in the classroom, having both planned and unplanned visits by the evaluators, asking more evaluators to visit classrooms, and engaging in pre- and post-conferences to discuss the collected data. In order for an evaluator to provide accurate feedback, they must spend adequate time in the classroom observing teachers (Resmovits, 2013). By setting up different types of visits, including unplanned mini-observations, the evaluator collects an array of snapshots that help increase the accuracy of the collected data (Marshall, 2005). Multiple evaluators, with various content knowledge and expertise,
provide diverse feedback to teachers (Donaldson, 2010). By scheduling pre- and post-conferences, teachers and evaluators develop stronger relationships. They also engage in reflective dialogue and strong coaching moments (Goldhammer, 1969; Marshall, 2014a).

The participants reported the need for evaluators to spend more time in the classrooms. They recognize that evaluators are busy. However, in order to collect authentic data and truly understand the complexities of each classroom, the evaluators needed to observe each class multiple times. Principals need to frequently visit classrooms and provide prompt feedback (Mendels, 2012). Some evaluators stopped in for only portions of the class periods, yet the evaluators scored the teachers in a variety of areas. In order to provide scores in multiple areas, it takes time to observe all of the components included in the evaluation system. The evaluator must observe multiple interactions between the students and teacher in order to understand the challenges faced in that specific environment (Resmovits, 2013). Evaluators must engage in frequent and focused walkthroughs in order to better understand the patterns in curriculum and instruction used by the teachers and provide accurate feedback (Downey et al., 2004; Resmovits, 2013; Wright, Horn, & Sanders, 1997). With frequent visits, the observations become less like “dog and pony shows” and a more representative of what occurs on a daily basis. The participants also recommended that evaluators engage in both planned and unplanned visits to the classrooms. By seeing teachers and students on multiple occasions, the evaluators more fully grasped the data that they collect and how to best help the teacher reflect, grow, and flourish.

The participants stated that having more than one person observe their classrooms would greatly benefit the teacher evaluation system (Donaldson, 2010). By having more people observe a teacher with their students, the teacher will receive additional data from...
multiple perspectives that focus on different aspects of the classroom environment. The teacher can discover patterns and truly reflect on how to improve their practices.

In order to make sure that the evaluator understands all areas of the classroom and to allow for professional growth, the participants commented that pre- and post-conferences must be prioritized (Marshall, 2014a). It is important that the evaluator and teacher discuss the dynamics of the classroom before the observation, and review the collected data after the observation, especially when teachers are working with various and challenging subgroups of students. The evaluation fails to differentiate for teachers’ professional needs (Weisberg et al., 2009). These conferences provide teachers and evaluators with the opportunity to discuss the professional needs of the teacher and ask questions specifically about the observations. It also allows the teacher an opportunity to provide any additional information that was missed during the observation. This process will increase the authenticity of the collected data.

In summary, the participants expressed concerns about the accuracy of the data collected during observations. Participants recommended that implementing additional methods could improve accuracy. The suggested methods include increasing the time evaluators spend in the classrooms, engaging in planned and unplanned classroom visits, including additional evaluators when observing classrooms and collecting data, and engaging in pre- and post-conferences. Multiple observations need to occur in order for evaluators to collect accurate observation data (Resmovits, 2013; Wright, Horn, & Sanders, 1997).

**Evaluators Need to Know the Students**

The participants stated that the evaluators need to know the students that they are observing. Various subgroups of students have different needs. In order for the evaluator to collect accurate data, they must understand the specific challenges faced by the students and
how those challenges impact the classroom environment and the strategies used by the teacher. The evaluators’ data should reflect the differences found in multiple classrooms. Some teachers work with students who live in socioeconomically challenged areas with very little parent support in the home. These students face different day-to-day challenges than students with more involved parents. There are students who are food insecure or experiencing hunger, and not having their basic needs met in the home. Some students are still learning English, and others have physical, emotional, behavioral, or learning challenges. The various dynamics that these challenges introduce into the classroom need to be reflected in the data collected during the observations. In order for this to occur, the evaluators must know the students in each classroom. Teacher evaluations must cater to the professional needs of the teachers, in order to support the students (Weisburg et al., 2009).

The most important component of the teacher evaluation is enhancing student achievement (Tucker & Strong, 2005). The current evaluation focuses on all students achieving grade-level performance. While that is an important level of performance to measure, improvements in student skills should also be recognized. The participants stated that the evaluation needs to stress student growth. If a teacher helps a student gain two grade levels of growth, yet the student has still not achieved grade-level status, the teacher’s evaluation is negatively impacted. The participants expressed frustration with the way teacher evaluations are negatively impacted by students who have made incredible gains, yet are still not on grade-level. Again, if the evaluators understood the needs of the students, they could better understand the growth achieved and the impact the teachers were making. As a result, a few participants mentioned potentially moving to schools that had more parent
involvement and less challenging student populations in order to decrease the stress of the evaluation process.

In summary, participants discussed the need for evaluators to know their students in order to understand the specific challenges faced in the classrooms and collect accurate data.

Conclusion

When asked about their perspectives on the observation portion of the teacher evaluation system, teachers focused on five key areas. These areas included evaluator and teacher relationships, frustration felt due to an unclear objective, lack of accuracy in the data, and needing evaluators to know the students and their individual needs. All five of the participants stated that their relationships with evaluators either positively or negatively impacted their observation data, in their perspective. They also indicated that the objective surrounding the observation process seemed unclear, which led to frustration. These five teachers stated that the observation data should focus on helping teachers grow, however, it appeared that the data was used to get rid of teachers. The participants also questioned the accuracy of the observation data. Because evaluators spent limited time in their classrooms, they questioned the data. They all felt that evaluators should come for planned and unplanned visits and more than one evaluator should visit each classroom to provide additional data. From the perspective of the four participants, evaluators also needed to understand that different groups of students have various needs. Before evaluators collect data, they should engage in a conference with the teacher to get a better understanding about the dynamics of the classroom and individual needs of the students. From the perspectives of the participants, changes need to occur in order to increase the effectiveness of the observation process.
Chapter 11: Recommendations

The recommendations listed below stem from themes that emerged during the study. The recommendations included emphasize a need for future research on actions that could be taken by school administrators and district personnel to improve the process of observation, as well as areas where teachers can engage in the process to improve their experiences with the observation portion of the evaluation system.

Recommendations for School Administrators and School District Personnel

Based on the findings from this research, school administrators and school district personnel should take into account the perspectives of teachers on the observation portion of the evaluation system, in order to make the process more effective. Teachers recognize the need for observations. However, they need to believe that the process is equitable and helps them grow as professionals. In order to make improvements to the process, the teachers need assistance from administrators and school district personnel.

For evaluators to build relationships with teachers and get to know the students in the teachers’ classrooms, they need to spend more time in the classrooms. Observations need to include both short and long visits, planned and unplanned. By increasing the time spent in the classrooms, evaluators collect more data and authentically understand the dynamics of the classroom (Donaldson, 2010). The accuracy of the data will increase, helping teachers reflect and improve their practices. Evaluators often face urgent situations during school days, which take away from their time spent in classrooms, but the observations must be prioritized (Ginsberg & Murphy, 2002).

Teachers and evaluators need to consistently meet for pre- and post-conferences in order to talk about the classroom environment and instructional strategies before the
observation, as well as confirm the data after the observations. When these discussions occur, the accuracy of the data improves and the teachers and evaluators build relationships through the interactions. Also, as a result of the conferences, both parties have an opportunity to talk about the students in the classroom to help the evaluator get to know the students better.

Administrators and school district personnel should also consider having more than one evaluator observe teachers’ classrooms (Donaldson, 2010). Teachers will benefit from receiving data and feedback from multiple perspectives. Scores from multiple evaluators will also help increase the accuracy of the data and diminish the impact for teachers who do not have a positive relationship with one of the evaluators.

Administrators and teachers should receive the same, consistent training that focuses on the objectives of the teacher evaluation system and its proper implementation. If all parties receive the same training together, they will be working off of the same information, and better understand the objectives and each other’s roles in the process. The lack of clarity about the objectives and the implementation of the teacher evaluation system lead teachers to feel frustrated.

In summary, from the findings of this research, the suggestions to administrators and school district personnel include requiring evaluators to spend more time in classrooms, visit the classrooms for planned and unplanned observations, and incorporate both pre- and post-conferences into the evaluation system. Administrators and school district personnel should require multiple evaluators to observe each classroom. Also, administrators and teachers should both engage in the same training to better understand the objectives of the teacher evaluation system and the implementation process. Teachers’ negative emotions will
decrease because they will understand the purpose and goals more clearly, and the system will become overall more effective and accurate. By executing these recommendations, teachers will build stronger relationships with evaluators, evaluators will get to know the students better, the accuracy of the data and scores will increase, and teachers overall negative reactions of the evaluation system will diminish.

**Recommendations for Teachers**

Based on themes that emerged through the interviews, the primary recommendation for teachers is more communication with their evaluators overall, and initiated by the teachers instead of the expectation that all of the communication should come from evaluators. If teachers believe that pre- and post-conferences are needed to improve the accuracy of the data and scores, and these conferences are not consistently used at their schools, they should schedule times with the evaluators before and after the observation sessions. By setting up these meetings, the teachers are proactively working on building collaborative relationships with the evaluators, sharing additional information about the classroom environment and the students, and finally reviewing the collected data and checking for accuracy. By increasing communication with the evaluators, teachers will better understand the goals of the evaluation and can help improve the overall accuracy and authenticity of the evaluation system. Being proactive about the communication and having an opportunity to share information about their strategies and students could also help empower teachers and reduce the emotional reaction teachers have toward the evaluation system.

Another recommendation for teachers is to invite more people to observe in their classrooms, in order to receive additional feedback. If the current evaluation system requires
only one evaluator to observe in each classroom, and a teacher wants additional feedback about his practices, he should invite teachers and administrators to observe his classroom. Additional data from multiple perspectives can help teachers better reflect on and understand their own practices and which areas might need improvement. Different educators will notice various aspects of the classroom environment and bring different perspectives and ideas for professional development. Also, if teachers believe that evaluators are not spending enough time in their classrooms, they can also invite evaluators visit more often.

Teachers should actively engage in the same training received by administrators. Both parties need to attend district-level training that is focused on the objectives of the teacher evaluation system, how to best implement the system, and the roles and responsibilities of the administrators and teachers. By completing the same training together, the process is more transparent and both parties will be working off of the same information. This clarity will decrease teacher frustration and increase overall understanding of the process for all parties.

**Recommendations for Future Research**

As this study encompassed only teachers in one city, future research should be conducted on teachers’ perspectives in other areas of the state about the observation process within the evaluation system. The data collected during the interviews of this study indicated the following: a) Evaluators need to spend more time in classrooms, including both planned and unplanned visits; b) Classrooms need to be visited by more than one evaluator to provide additional data to teachers; and c) Evaluators and teachers consistently need to engage in pre- and post-conference meetings.
The study participants stated that the purpose of the classroom observations would become clearer and the overall accuracy of data and scores would increase if these additional recommendations were followed. By conducting future research with participants in other areas of the state, it will help to see if teacher responses in other areas correlate to the responses of the participants of this study. Involving additional participants in the current state and in other states across the country, to broaden the research, would provide additional data and insights into teacher evaluation processes and classroom observation protocols. This data would help to discover the extent to which the findings of this study can be generalized. Research explains the importance of observations, but it is challenging to assess the perceptions of teachers and to understand how those perceptions impact the overall evaluation system.

The findings reflect a limited glance into teacher perspectives of the observation portion of the evaluation system. Until additional research is conducted, the recommendations that extend from the findings should be regarded as suggestions rather than conclusions. At this point, the findings are strong enough to offer some recommendations for further discussion and research to school administrators, district personnel, and teachers.

Summary

The qualitative data collected and analyzed during my phenomenological study reflected teachers’ perceptions of the observation portion of the evaluation system. All five of the participants recognized the need for incorporating observations into the evaluation system. However, the participants believed that changes needed to occur in order to increase the effectiveness and accuracy of the process. The changes included having evaluators spend more time in the classrooms, more evaluators visiting each classroom, and pre- and post-
conferences. Also, teachers and evaluators should engage in training together to better understand the teacher evaluation system. These adjustments could help teachers build stronger relationships with their evaluators and increase overall accuracy of collected data, as well as help evaluators know the students better. By building stronger relationships between the teachers, evaluators, and students, the school culture would be positively impacted, and overall productivity and commitment towards achieving core values and goals would increase. If these changes occurred, the participants indicated that the process would be more meaningful and helpful as they aim to strengthen their practices, which ultimately benefits the students and their achievement.

Social constructivism guides the implementation of the recommendations. All of the recommendations entail social and interactive processes and connecting knowledge to personal experience. These social interactions and dialogue, including conferences, additional classrooms visits, and observations by multiple evaluators, help create knowledge. Teachers and evaluators need additional collaboration and interactions to construct understanding while also integrating their own life experiences and culture. Through these social interactions and dialogues, teachers and administrators will continue to understand best practices and how to integrate them in the classroom to positively impact student learning. Also, teachers and administrators engage in analyzing their own practices, connecting their findings to prior experiences, and continue to grow through the process (Dewey, 1938). These recommendations center on how interactions and dialogues between teachers and evaluators can help decrease frustrations and impact their overall perspectives of the accuracy and importance of the teacher evaluation process (Deal & Peterson, 2011).
Appendices

Appendix A: Interview Questions ................................................................. 160
Appendix B: Observation Rubric................................................................. 161
Appendix C: Domains.................................................................................. 162
Appendix A: Interview Questions

1. What has been your overall teacher evaluation experience throughout your career?

2. What has been your overall experience with the observation process within the evaluation system?

3. What is the purpose of the current evaluation process from your perspective?

4. In this southwestern state, how effective is the current observation instrument from your perspective?

5. What components of the observation process are most useful, from your perspective?

6. What components of the observation process are not useful, from your perspective?

7. In your experience, what data is collected during the observation process?

8. From your perspective, how useful is the data collected during the observation process?

9. From your perspective, what changes, if any, should be made to improve the current observation system?

10. What other thoughts or comments would you like to share about the current observation process?
## Appendix B: Observation Rubric

<table>
<thead>
<tr>
<th>Elements</th>
<th>Strands</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A: Demonstrating knowledge of content</td>
<td>Knowledge of Content and Pedagogy</td>
<td>Preparation for Learning</td>
</tr>
<tr>
<td>1B: Designing Coherent Instruction</td>
<td>Knowledge of Students</td>
<td>Domain 1: Planning and Organizing</td>
</tr>
<tr>
<td>1C: Setting instructional outcomes</td>
<td>Creating an Environment of Respect &amp; Rapport</td>
<td>Domain 2: Teaching and Assessing</td>
</tr>
<tr>
<td>1D: Demonstrating knowledge of resources</td>
<td>Establishing a Culture of Learning</td>
<td>Domain 3: Communicating and Collaborating</td>
</tr>
<tr>
<td>1E: Designing Student Assessment</td>
<td>Managing Classroom Procedures</td>
<td>Domain 4: Professional Growth</td>
</tr>
<tr>
<td>1F: Creating an environment of respect and report</td>
<td>Managing Student Behavior</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Domains

### Domain 1: Preparation and Planning

#### Element: NMTEACH 1A: Demonstrating knowledge of content
- To what level is content communicated in the lesson plan and resulting lesson?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>Teacher’s plans display little knowledge of the content and no alignment to NM adopted standards.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Teacher’s plans reflect some knowledge of the content and partial alignment to NM adopted standards.</td>
</tr>
<tr>
<td>Effective</td>
<td>Teacher’s plans reflect solid knowledge of the content and are clearly aligned to NM adopted standards. Teacher demonstrates familiarity with resources to enhance own knowledge in each core area.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Teacher’s plans reflect extensive knowledge of the content in core areas. Teacher’s instructional plans incorporate research and resources related to the NM adopted standards.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>Teacher’s plans reflect extensive knowledge of content. Teacher incorporates current research resources to support NM adopted standards. Teacher contributes to the refinement and development of the approved NM adopted standards-aligned curriculum.</td>
</tr>
</tbody>
</table>

#### Notes:

#### Element: NMTEACH 1B: Designing coherent instruction
- To what level are activities meaningfully sequenced to support learning?
- To what level are a variety of learning strategies used within the instructional plan?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>The sequence of learning experiences is poorly aligned with NM adopted standards.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>The sequence of learning experiences demonstrates partial alignment with NM adopted standards.</td>
</tr>
<tr>
<td>Effective</td>
<td>The lesson is designed to implement instructional targets aligned to NM adopted standards as follows: Creating explicit connections between previous learning and new concepts and skills; contains substantive learning tasks; structure learning tasks progressively to develop students' cognitive abilities and skills. The sequence of learning experiences is aligned to NM adopted standards, instructional learning targets and is differentiated by scaffolding content and academic language for diverse learners.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Teacher designs pedagogical practices, including student grouping, differentiated instruction based on student level, and prepared questions to reinforce and extend student learning to include real world, application-based experiences.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>The teacher shows evidence of designing coherent instruction in a collaborative manner by intentionally demonstrating awareness and processes for engaging all students.</td>
</tr>
</tbody>
</table>

#### Notes:
### Domain 1: Preparation and Planning

#### Element: NMTEACH 1C: Setting instructional outcomes
- How are daily learning goals communicated to students?
- To what level do learning goals directly align to content standards?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>Instructional targets are not aligned to NM adopted standards.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Instructional targets are moderately aligned to NM adopted standards, but not explicitly stated to students.</td>
</tr>
<tr>
<td>Effective</td>
<td>Instructional targets are aligned to NM adopted standards and stated as measurable and observable goals for student learning. Instructional processes and activities address students' varying abilities, and are aligned to instructional targets.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Instructional targets are aligned to NM adopted content, and are translated into student accessible learning objectives. The instructional process and learning activities are rigorous and aligned to NM adopted standards and instructional outcomes, and include plans for modifications to ensure students are able to complete the targeted objective.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>The teacher has a deep understanding of grade-level NM adopted standards and appropriate pedagogy to ensure all students are making progress toward deep understand and proficiency in NM adopted standards and learning targets.</td>
</tr>
</tbody>
</table>

#### Element: NMTEACH 1D: Demonstrating knowledge of resources
- How does the teacher utilize skills and content learned from professional development opportunities?
- What resources have been provided to students to support learning?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>Teacher demonstrates little or no familiarity with resources to enhance own content knowledge, to use in teaching, or for students who demonstrate need.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Teacher demonstrates some familiarity with resources to enhance own content knowledge, to use in teaching, or for students who demonstrate need.</td>
</tr>
<tr>
<td>Effective</td>
<td>Teacher fully utilizes existing resources, including support materials, textbooks, supplementary materials, to enhance content knowledge, to use in teaching, or for students who demonstrate need.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Teacher seeks out and uses resources beyond school/district, in professional organizations, internet, and community to enhance content knowledge, to use in teaching, or for students who demonstrate need.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>The teacher actively engages colleagues and provides resources to them in areas that are pertinent to their needs. The teacher also collects and shares content specific research studies and practices, and shares outside resources. Teacher provides and trains staff for school-wide initiatives.</td>
</tr>
</tbody>
</table>

---
### Domain 1: Preparation and Planning

#### Element: NMTEACH 1E: Demonstrating knowledge of students
- To what level have student learning styles been addressed in the lesson?
- How has student achievement data been used to design activities to support content acquisition?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>Teacher demonstrates little or no knowledge of students' backgrounds, cultures, skills, academic language development, interests, and special needs.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Teacher demonstrates some knowledge of students' backgrounds, cultures, skills, academic language development, interests, and special needs.</td>
</tr>
<tr>
<td>Effective</td>
<td>Teacher demonstrates solid knowledge of students' backgrounds, cultures, skills, academic language development, interests, and special needs.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Teacher demonstrates extensive knowledge of students' backgrounds, cultures, skills, academic language development, interests, and special needs, and incorporates culturally-sensitive strategies into instructional planning and practice.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>The teacher provides novice and struggling teachers with understanding, resources, and mentorship for addressing the unique needs of individual students. The teacher provides ongoing support to administration in demonstrating linguistically- and culturally-appropriate instructional programs for the school site.</td>
</tr>
</tbody>
</table>

#### Notes:

---

### Domain 1: Preparation and Planning

#### Element: NMTEACH 1F: Designing student assessment
- To what level has the teacher incorporated formative assessment techniques throughout the lesson?
- How are students assessed to determine understanding of the learning target at the end of the lesson?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>Teacher's plan for assessing student learning contains no clear criteria or NIM adopted standards, is poorly aligned with the instructional outcomes, or is inappropriate to many students.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Teacher's plan for assessing student learning is partially aligned with the instructional outcomes, and is appropriate to some students.</td>
</tr>
<tr>
<td>Effective</td>
<td>Teacher's plan for assessing student learning is aligned with the instructional outcomes, success criteria and the assessment tools. Teacher uses clear criteria to produce evidence which enables the teacher to make instructional adjustments and provide feedback to move student learning forward. Formative assessments are explicitly planned for each incremental learning step to ensure student learning outcomes.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>There is full alignment between the instructional outcomes, the success criteria, and the assessment tools. Teacher uses clear criteria that show where each student is in his/her learning.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>The teacher helps implement school-wide training and implementation for understanding and using assessment data. Students are assessed in multiple ways, using a variety of approaches to show what they know and where they are in their learning.</td>
</tr>
</tbody>
</table>

#### Notes:
## Domain 2: Creating an Environment for Learning

<table>
<thead>
<tr>
<th>Element:</th>
<th>NMTEACH.2A: Creating an environment of respect and rapport</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>- To what level are interactions in the classroom positive and productive?</td>
</tr>
<tr>
<td></td>
<td>- To what level are all student groups respected and valued in the classroom?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classroom interaction both between the teacher and students, and among students, are inappropriate or insensitive to students' cultural backgrounds, and may include the following:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Sarcasm.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Put-downs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Conflict.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Classroom interactions, both between the teacher and students, and among students, are generally positive, but may include these:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Some conflict.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Occasional displays of insensitivity.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Occasional lack of responsiveness to cultural or developmental differences among students.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Classroom interactions, between teacher and students, and among students, are as follows:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Are polite and respectful.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Demonstrate knowledge of cultural and developmental differences among groups of students.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Disagreements are handled respectfully.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Classroom interactions among the teacher and individual students are as follows:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Are highly respectful.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Reflect warmth and caring.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Practice reflects sensitivity to students' cultures and levels of development.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Respectful discourse.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>In addition to all the requirements to be highly effective, the teacher as a leader demonstrates the following:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Helps create a school-wide environment of respect for the campus, the stakeholders, and the rules.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Works with colleagues on developing support for students in need.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Helps to create school-wide interventions, and support programs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
## Domain 2: Creating an Environment for Learning

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Element:</strong></td>
<td>NMTEACH 2B: Organizing physical space</td>
<td>The physical environment is as follows:</td>
<td>The classroom is safe as follows:</td>
<td>The classroom is safe as follows:</td>
<td>In addition to all the requirements to be highly effective, the teacher as a leader does the following:</td>
</tr>
<tr>
<td><strong>Ineffective</strong></td>
<td>To what level do all students have equal access to learning resources and materials?</td>
<td>• Unsafe.</td>
<td>• Learning is accessible to all students.</td>
<td>• Students contribute to the use or adaptation of the physical environment to advance learning.</td>
<td>• Teacher uses the classroom to model or demonstrate for other teachers.</td>
</tr>
<tr>
<td></td>
<td>To what level does the classroom environment support the day’s lesson?</td>
<td>• Students do not have access to learning.</td>
<td>• Teacher ensures that the physical arrangement is appropriate to the learning activities.</td>
<td>• Technology is used skillfully, by teachers as appropriate to the lesson.</td>
<td>• Helps colleagues arrange their environment so learning is accessible to all.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Poor alignment between the environment and the lesson activities.</td>
<td>• There is posted evidence of student learning.</td>
<td>• Technology is used skillfully, by teachers and students as appropriate to the lesson.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Teacher is partially effective in modifying the environment to suit learning activities.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
<table>
<thead>
<tr>
<th>Element:</th>
<th>Domain 2: Creating an Environment for Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NMTEACH 2C: Establishing a culture for learning</td>
</tr>
<tr>
<td></td>
<td>• To what level do students exhibit a learning energy during the lesson that supports engagement?</td>
</tr>
<tr>
<td></td>
<td>• To what level are students encouraged to communicate with others to address learning goals?</td>
</tr>
<tr>
<td>Ineffective</td>
<td>The classroom environment conveys a negative culture for learning as follows</td>
</tr>
<tr>
<td></td>
<td>• Low teacher commitment to the subject.</td>
</tr>
<tr>
<td></td>
<td>• Low expectations for student achievement.</td>
</tr>
<tr>
<td></td>
<td>• Little or no student effort.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Attempts to create a culture for learning and is partially successful as follows:</td>
</tr>
<tr>
<td></td>
<td>• Some teacher commitment to the subject.</td>
</tr>
<tr>
<td></td>
<td>• Modest expectations for student achievement.</td>
</tr>
<tr>
<td></td>
<td>• Some student effort.</td>
</tr>
<tr>
<td></td>
<td>• Teacher and students appear to be “going through the motions.”</td>
</tr>
<tr>
<td>Effective</td>
<td>The classroom culture is characterized by high expectations for all students.</td>
</tr>
<tr>
<td></td>
<td>• The teacher establishes norms and participant structures in which students can learn with and from each other, i.e. student grouping, student presentations, and peer editing.</td>
</tr>
<tr>
<td></td>
<td>• Teacher conveys content relevance.</td>
</tr>
<tr>
<td></td>
<td>• Demonstrated commitment to the subject by both teacher and students.</td>
</tr>
<tr>
<td></td>
<td>• Students demonstrate pride in their efforts.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Culture for learning in which everyone shares a belief in the importance of the subject as follows:</td>
</tr>
<tr>
<td></td>
<td>• High levels of student excitement and teacher passion for the subject.</td>
</tr>
<tr>
<td></td>
<td>• Students hold themselves to high standards of performance Students initiate improvements to their efforts.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>In addition to all the requirements to be highly effective, the teacher as a leader does the following:</td>
</tr>
<tr>
<td></td>
<td>• Promotes and organizes school-wide learning program(s) and learning culture among all stakeholders.</td>
</tr>
<tr>
<td>Notes:</td>
<td></td>
</tr>
</tbody>
</table>


## Domain 2: Creating an Environment for Learning

<table>
<thead>
<tr>
<th>Element: NMTEACH 2D: Managing classroom procedures</th>
</tr>
</thead>
<tbody>
<tr>
<td>• To what level is the classroom culture and routine maximizing instructional time?</td>
</tr>
<tr>
<td>• To what level does the teacher use developmentally appropriate procedures to maximize instructional time.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Notes:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ineffective</strong></td>
<td>Instructional time is lost.</td>
</tr>
<tr>
<td></td>
<td>• Inefficient classroom routines.</td>
</tr>
<tr>
<td></td>
<td>• Inefficient procedures for transition.</td>
</tr>
<tr>
<td></td>
<td>• Inefficient use of supplies.</td>
</tr>
<tr>
<td><strong>Minimally Effective</strong></td>
<td>Some instructional time is lost.</td>
</tr>
<tr>
<td></td>
<td>• Partially-effective classroom routines and procedures</td>
</tr>
<tr>
<td></td>
<td>• Partially-effective routines for transition</td>
</tr>
<tr>
<td></td>
<td>• Partially-effective use of supplies</td>
</tr>
<tr>
<td><strong>Effective</strong></td>
<td>Little instructional time is lost.</td>
</tr>
<tr>
<td></td>
<td>• Effective classroom routines and procedures.</td>
</tr>
<tr>
<td></td>
<td>• Teacher leads effective routines for transition.</td>
</tr>
<tr>
<td></td>
<td>• Effective use of supplies.</td>
</tr>
<tr>
<td><strong>Highly Effective</strong></td>
<td>Students contribute to the seamless operation of the classroom.</td>
</tr>
<tr>
<td></td>
<td>• Routines and procedures are evident.</td>
</tr>
<tr>
<td></td>
<td>• Effective transitions and use of supplies.</td>
</tr>
<tr>
<td></td>
<td>• Students lead effective routines for transition.</td>
</tr>
<tr>
<td><strong>Exemplary</strong></td>
<td>In addition to all the requirements to be highly effective, the teacher as a leader helps to create a culture of student ownership of school-wide operations.</td>
</tr>
</tbody>
</table>
### Domain 2: Creating an Environment for Learning

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element:</td>
<td>NMTEACH 2E: Managing student behavior</td>
<td>No evidence that standards of conduct have been established.</td>
<td>Teacher has made an effort to establish standards of conduct for students</td>
<td>Standards of conduct are designed to create an atmosphere conducive to learning, with a focus on self-discipline, respecting the rights of others, and cooperating with one another.</td>
<td>In addition to standards being clear to students are these elements:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Little or no teacher monitoring of student behavior.</td>
<td>• Effort made with inconsistent results to monitor students’ behavior.</td>
<td>• Standards are clear to students.</td>
<td>• Evidence of student participation in setting conduct standards.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Response to student misbehavior is regressive or disrespectful of student dignity.</td>
<td>• Response to student misbehavior is inconsistent.</td>
<td>• Teacher holds students responsible for maintaining behavioral standards.</td>
<td>• Teacher’s monitoring of student behavior is highly effective.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Teacher response to student misbehavior is appropriate and respects the students’ dignity.</td>
<td>• Teacher’s response to student misbehavior is sensitive to individual needs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Teacher response is consistent.</td>
<td>• Students take an active role in monitoring the standards of behavior.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>In addition to all the requirements to be highly effective, the teacher as a leader demonstrates the following:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Actively engages in the monitoring of student behavior school-wide.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Serves as a model of positive behavior for stakeholders.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Teacher promotes system(s) of school-wide positive behavioral support that encourages stakeholders to promote and monitor a safe and healthy environment.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Notes:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Domain 3: Teaching for Learning

#### Element: NMTEACH 3A: Communicating with students in a manner that is appropriate to their culture and level of development
- To what level are directions clearly delivered and understandable?
- To what level is content communicated in a clear, concise manner?

| Level of Performance | | |
|----------------------|-------------------------------|
| Ineffective | Does not deliver clear expectations for learning, directions, procedures, and explanations of content to students |
| Minimally Effective | Limited expectation for learning, directions, procedures, and explanation of content |
| Effective | Teacher uses clear communication employing a range of vocabulary to ensure learning expectations are comprehensible to all students. Teacher allows for student clarification and feedback. |
| Highly Effective | Expectation for learning, directions, procedures, and explanation of content are evident, consistent, and anticipate possible student misconceptions |
| Exemplary | The highly effective teacher promotes ongoing and consistent communication with students. Students are provided multiple opportunities and/or modalities to express concepts being taught in class and are clearly aware of their progress with those concepts |

#### Notes:

### Domain 3: Teaching for Learning

#### Element: NMTEACH 3B: Using questioning and discussion techniques to support classroom discourse
- To what level do all students have an opportunity to answer questions?
- To what level are questions thought provoking and rigorous?

| Level of Performance | | |
|----------------------|-------------------------------|
| Ineffective | Teacher questioning techniques are not aligned to content and provide no opportunity for student engagement |
| Minimally Effective | Teacher questioning techniques are low-level with minimal student engagement |
| Effective | The teacher’s questioning techniques elicit a deep response and allows for sufficient time for students to answer through active engagement with peers and teacher |
| Highly Effective | The teacher promotes consistent analytical and collaborative approaches to understanding, uses questioning techniques that scaffold instruction for deep understanding of concepts, allowing for discussion and debate of key concepts |
| Exemplary | Questioning techniques are engaging and reflect a high level of thinking in a culturally and developmentally appropriate environment. Students engage in deep meaningful conversations using academic language |

#### Notes:
## Domain 3: Teaching for Learning

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Element: NMTEACH 3C: Engaging students in learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>To what level are students engaging in the lesson’s activities?</td>
</tr>
<tr>
<td></td>
<td>To what level are activities sequential and aligned to the daily learning target?</td>
</tr>
<tr>
<td></td>
<td>To what level are students required to be intellectually engaged with the course content?</td>
</tr>
<tr>
<td></td>
<td><strong>Ineffective</strong></td>
</tr>
<tr>
<td></td>
<td>Activities, assignments, materials, and grouping of students are inappropriate to the instructional outcomes, resulting in no intellectual engagement.</td>
</tr>
<tr>
<td></td>
<td>The lesson has no structure and/or is poorly paced.</td>
</tr>
<tr>
<td></td>
<td><strong>Minimally Effective</strong></td>
</tr>
<tr>
<td></td>
<td>Activities, assignments, materials, and grouping of students are somewhat appropriate to the instructional outcomes, resulting in moderate intellectual engagement.</td>
</tr>
<tr>
<td></td>
<td>The lesson does not connect to prior understanding.</td>
</tr>
<tr>
<td></td>
<td>The lesson has a recognizable structure, but is not fully maintained.</td>
</tr>
<tr>
<td></td>
<td>The lesson does not have clear learning goals (more specific than broad standard).</td>
</tr>
<tr>
<td></td>
<td><strong>Effective</strong></td>
</tr>
<tr>
<td></td>
<td>Activities, assignments, materials, and grouping of students are fully appropriate to the instructional outcomes.</td>
</tr>
<tr>
<td></td>
<td>The lesson explicitly connects to prior understanding.</td>
</tr>
<tr>
<td></td>
<td>All students are engaged.</td>
</tr>
<tr>
<td></td>
<td>The lesson’s structure is coherent and paced appropriately.</td>
</tr>
<tr>
<td></td>
<td>The lesson has specific learning goals aligned to the standard.</td>
</tr>
<tr>
<td></td>
<td>The lesson allows for student reflection.</td>
</tr>
<tr>
<td></td>
<td><strong>Highly Effective</strong></td>
</tr>
<tr>
<td></td>
<td>Activities, assignments, materials, and grouping of students are designed to support challenging instructional outcomes.</td>
</tr>
<tr>
<td></td>
<td>Students are highly intellectually engaged.</td>
</tr>
<tr>
<td></td>
<td>The lesson is adapted as needed to the readiness of each student and the structure and pacing allow for students’ reflection and closure.</td>
</tr>
<tr>
<td></td>
<td>The lesson allows for formative assessment.</td>
</tr>
<tr>
<td></td>
<td><strong>Exemplary</strong></td>
</tr>
<tr>
<td></td>
<td>Expectations of students are at an advanced level to engage learners to obtain depth of knowledge.</td>
</tr>
<tr>
<td></td>
<td>The teacher formatively assesses student engagement, understanding, and ability to analyze, and immediately adapts methods for improved learning.</td>
</tr>
</tbody>
</table>

**Notes:**
### Domain 3: Teaching for Learning

#### Element: NMTEACH 3D: Assessment in Instruction
- To what level does the teacher determine the understanding and needs of each student during the lesson?
- To what level are students aware of how they will demonstrate understanding of the content/lesson?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessments are not used in instruction.</td>
<td>Assessments are minimally used in instruction.</td>
<td>Assessments are consistently used in instruction.</td>
<td>Assessments are used in a sophisticated manner to drive instruction:</td>
<td>Assessments are used in a sophisticated manner to drive instruction:</td>
<td>Students analyze and evaluate assessment data, and information, and apply same to improved learning.</td>
</tr>
<tr>
<td>Students are unaware of assessment criteria.</td>
<td>Students are minimally aware of the assessment criteria.</td>
<td>There are clear goals and performance criteria, communicated effectively to students.</td>
<td>The teacher establishes, supports, and models the use of consistent assessment of progression and development as a tool for improved learning to stakeholders.</td>
<td>The teacher establishes, supports, and models the use of consistent assessment of progression and development as a tool for improved learning to stakeholders.</td>
<td>The teacher involves students in establishing the assessment criteria and provides high quality feedback from a variety of sources.</td>
</tr>
<tr>
<td>The teacher does not monitor student progress or offer feedback.</td>
<td>The teacher occasionally monitors students’ progress and provides limited or irrelevant feedback.</td>
<td>The assessment strategies are aligned to the goal and criteria, and elicit evidence during instruction.</td>
<td>The teacher involves students in establishing the assessment criteria and provides high quality feedback from a variety of sources.</td>
<td>The teacher involves students in establishing the assessment criteria and provides high quality feedback from a variety of sources.</td>
<td>The teacher involves students in establishing the assessment criteria and provides high quality feedback from a variety of sources.</td>
</tr>
</tbody>
</table>

#### Element: NMTEACH 3E: Demonstrating flexibility and responsiveness
- To what level does the teacher modify instruction within the lesson/class period?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher adheres to the instructional plan, even when a change would maximize learning.</td>
<td>Teacher accepts responsibility for student success.</td>
<td>Teacher promotes the successful learning of all students.</td>
<td>Teacher identifies unique “teachable moments” that relate current lessons/standards to individual and student groups.</td>
<td>The teacher identifies unique “teachable moments” that relate current lessons/standards to individual and student groups.</td>
<td>The teacher identifies unique “teachable moments” that relate current lessons/standards to individual and student groups.</td>
</tr>
<tr>
<td>The teacher disregards students’ learning challenges.</td>
<td>Teacher attempts to modify the lesson and responds to student questions with moderate success, but has a limited repertoire of strategies to draw upon.</td>
<td>The teacher adjusts instructional plans and makes accommodations for student questions, needs, and interests.</td>
<td>The teacher applies student interest to current learning goal.</td>
<td>The teacher applies student interest to current learning goal.</td>
<td>The teacher applies student interest to current learning goal.</td>
</tr>
<tr>
<td>The teacher blames the students or their environment for lack of academic progress.</td>
<td>The teacher utilizes a variety of strategies.</td>
<td>The teacher ensures the success of all students, using an extensive repertoire of instructional strategies.</td>
<td>The teacher identifies unique “teachable moments” that relate current lessons/standards to individual and student groups.</td>
<td>The teacher identifies unique “teachable moments” that relate current lessons/standards to individual and student groups.</td>
<td>The teacher identifies unique “teachable moments” that relate current lessons/standards to individual and student groups.</td>
</tr>
</tbody>
</table>
### Domain 4: Professionalism

**Element:** NMTEACH 4A: Communicating with families
- How well does the teacher engage families in the instructional program?
- To what level is the teacher’s communication (both formal and informal) with families frequent and culturally appropriate?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The teacher does not attempt to engage families in the instructional program.</td>
<td>The teacher makes minimal attempts to engage families in the instructional program.</td>
<td>The teacher successfully engages families in the instructional program.</td>
<td>The teacher successfully engages families in the instructional program.</td>
<td>The teacher helps promote school-wide activities that increase family and community understanding of the instructional program.</td>
</tr>
<tr>
<td></td>
<td><em>Teacher communication with families is sporadic or culturally inappropriate.</em></td>
<td><em>Teacher communication is not always appropriate to the cultures of families.</em></td>
<td><em>Teacher communicates with families in a culturally appropriate manner.</em></td>
<td><em>Teacher’s communications are sensitive to cultural traditions, and students participate in the communication.</em></td>
<td><em>Teacher helps promote school-wide activities that increase family involvement.</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>Teacher frequently communicates with families.</em></td>
<td><em>Teacher communicates frequently and effectively with families.</em></td>
<td><em>Teacher actively seeks out and engages with stakeholders within the community, and becomes a part of the community.</em></td>
</tr>
</tbody>
</table>

**Notes:**

---

**Element:** NMTEACH 4B: Participating in a professional community
- How willing and eager is the teacher to participate in the professional community?
- How collegial and productive are teacher’s relationships with their colleagues?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Ineffective</th>
<th>Minimally Effective</th>
<th>Effective</th>
<th>Highly Effective</th>
<th>Exemplary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The teacher does not participate in a professional community or in school and district events and projects.</td>
<td>The teacher participates in a professional community and in school and district events and projects when specifically requested.</td>
<td>The teacher participates actively in professional community, and in school/district events and projects.</td>
<td>The teacher makes a substantial contribution to the professional community, to school/district events and projects.</td>
<td>The teacher is actively engaging in ongoing research, leads study groups, and identifies new practices for school and district implementation.</td>
</tr>
<tr>
<td></td>
<td><em>Teacher’s relationships with colleagues are negative or self-serving.</em></td>
<td><em>Teacher’s relationships with colleagues are cordial but relationships do not lead to productive work that benefits students.</em></td>
<td><em>Teacher maintains positive and productive relationships with colleagues.</em></td>
<td><em>Teacher assumes a leadership role among the stakeholders.</em></td>
<td><em>Teacher serves as an instructional leader, and is accepted by faculty for exceptional skills in delivering professional development and mentorship.</em></td>
</tr>
</tbody>
</table>

**Notes:**
### Domain 4: Professionalism

#### Element: Reflecting on teaching

- **How detailed, accurate, and thoughtful is the teacher’s reflection on their instructional practices?**

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>Teacher does not accurately assess the effectiveness of the instructional practices. Teacher has no idea about how the instructional practices could be improved.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>Teacher provides a partially accurate and objective description of the instructional practices with some evidence. Teacher makes only general suggestions as to how the instructional practices might be improved.</td>
</tr>
<tr>
<td>Effective</td>
<td>Teacher provides an accurate and objective description of own and other instructional practices with specific evidence. Teacher makes some specific suggestions as to how the instructional practices might be improved.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>Teacher’s reflection on instructional practices is thoughtful and accurate with specific evidence. Teacher draws on an extensive repertoire to suggest alternative strategies and predicting the likely success of each.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>Teacher’s reflection is ongoing and immediate. The teacher demonstrates immediate understanding of effectiveness of instructional practices. Teacher modifies and adapts as necessary.</td>
</tr>
</tbody>
</table>

#### Element: Demonstrating professionalism

- **How high are the teacher’s professional standards and practices?**
- **To what level is the teacher willing to comply with district and school rules and regulations?**

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective</td>
<td>The teacher displays a lack of professionalism. Teacher contributes to practices that are self-serving or harmful to students. Teacher fails to comply with regulations and timelines.</td>
</tr>
<tr>
<td>Minimally Effective</td>
<td>The teacher displays minimal professionalism. Teacher complies inconsistently with regulations, doing just enough to “get by.”</td>
</tr>
<tr>
<td>Effective</td>
<td>The teacher displays a high level of professionalism in dealings with both students and colleagues. Teacher complies fully and voluntarily with regulations. Teacher promotes a safe environment for students when monitoring students and activities.</td>
</tr>
<tr>
<td>Highly Effective</td>
<td>The teacher is proactive and assumes a leadership role in ensuring the highest-level of professional practices by all colleagues. Teacher helps ensure that school practices honor all stakeholders. Teacher helps colleagues comply with rules and regulations.</td>
</tr>
<tr>
<td>Exemplary</td>
<td>The teacher proactively and in a positive way seeks to continually improve the culture of the school by consistently raising expectations for adults and students, raising the engagement of adults and students and contributing to the efficacy of adults and students.</td>
</tr>
</tbody>
</table>
### Domain 4: Professionalism

**Element:** NMTEACH 4E: Growing and developing professionally

- To what level does the teacher seek out, implement, and share professional learning?
- How well does the teacher utilize feedback?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ineffective</strong></td>
<td>The teacher does not participate in professional development activities.</td>
</tr>
<tr>
<td></td>
<td>- Teacher makes no effort to share knowledge with colleagues.</td>
</tr>
<tr>
<td></td>
<td>- Teacher is resistant to feedback from supervisors or colleagues.</td>
</tr>
<tr>
<td><strong>Minimally Effective</strong></td>
<td>The teacher participates in professional development activities that are convenient or are required.</td>
</tr>
<tr>
<td></td>
<td>- Teacher makes limited attempts to share knowledge with colleagues.</td>
</tr>
<tr>
<td></td>
<td>- Teacher accepts feedback from supervisors and colleagues with some reluctance.</td>
</tr>
<tr>
<td><strong>Effective</strong></td>
<td>The teacher accepts opportunities for professional development after an individual assessment of need.</td>
</tr>
<tr>
<td></td>
<td>- Teacher implements PD strategies.</td>
</tr>
<tr>
<td></td>
<td>- Teacher welcomes and implements feedback from supervisors and colleagues.</td>
</tr>
<tr>
<td><strong>Highly Effective</strong></td>
<td>The teacher actively pursues professional development opportunities.</td>
</tr>
<tr>
<td></td>
<td>- Teacher initiates activities to share expertise with others.</td>
</tr>
<tr>
<td></td>
<td>- Teacher seeks out feedback from supervisors and colleagues.</td>
</tr>
<tr>
<td><strong>Exemplary</strong></td>
<td>The teacher is an established leader in the school.</td>
</tr>
<tr>
<td></td>
<td>- Teacher is able to provide feedback to colleagues and supervisors in a manner that is welcomed and utilized by all stakeholders.</td>
</tr>
</tbody>
</table>

**Notes:**

---

### Domain 4: Professionalism

**Element:** NMTEACH 4F: Maintaining accurate records

- How efficient and accurate are the teacher’s record-keeping systems?

<table>
<thead>
<tr>
<th>Level of Performance</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ineffective</strong></td>
<td>The teacher’s systems for maintaining both instructional and non-instructional records are either non-existent or in disarray.</td>
</tr>
<tr>
<td></td>
<td>- Information from records contains errors and causes confusion.</td>
</tr>
<tr>
<td><strong>Minimally Effective</strong></td>
<td>The teacher’s systems for maintaining both instructional and non-instructional records are rudimentary and partially successful.</td>
</tr>
<tr>
<td></td>
<td>- Information from records is mostly accurate and not up to date.</td>
</tr>
<tr>
<td><strong>Effective</strong></td>
<td>The teacher’s systems for maintaining both instructional and non-instructional records are efficient and successful.</td>
</tr>
<tr>
<td></td>
<td>- Information from records is accurate and up to date.</td>
</tr>
<tr>
<td></td>
<td>- Information is used by teacher to make decisions regarding students.</td>
</tr>
<tr>
<td><strong>Highly Effective</strong></td>
<td>The students contribute to the maintenance of the efficient and successful systems for both instructional and non-instructional records.</td>
</tr>
<tr>
<td></td>
<td>- Information from records is accurate, up to date, and used constructively by students and teacher.</td>
</tr>
<tr>
<td><strong>Exemplary</strong></td>
<td>The teacher’s system is recognized by the school community as efficient and used as a model for other teachers.</td>
</tr>
</tbody>
</table>

**Notes:**


*Teaching and Teacher Education, 16*(5-6), 523-545.


Palincsar, A.S. (1998). *Social constructivist perspectives on teaching and learning.* Retrieved from: http://www.researchgate.net/publication/5285760_Social_constructivist_perspectives_on_teaching_and_learning?enrichId=rqreq-493df0d0-824d-418f-990a-bdb6ec919e18&enrichSource=Y292ZXJQYWdlOzUyODU3NjA7QVM6MTg3MzEzNzI0Mz M4MTCe5QDE0Mje2NzA0ODA5MTM%3D&el=1_x_2


Popham, W. J. (2007). Instructional insensitivity of tests: Accountability's dire drawback: If we plan to use tests for purposes of accountability, we need to know that they measure traits that can be influenced by instruction. *Phi Delta Kappan, 89*(2), 146.


