Fall 10-26-2016

An Exploration of Teacher Perception of the Marzano Causal Teacher Evaluation Model and Its Impact on Professional Practices

Susan Kathryn Graziano

Seton Hall University, graziasu@shu.edu

Follow this and additional works at: https://scholarship.shu.edu/dissertations

Part of the Educational Administration and Supervision Commons, and the Educational Assessment, Evaluation, and Research Commons

Recommended Citation

https://scholarship.shu.edu/dissertations/2234
AN EXPLORATION OF TEACHER PERCEPTION OF THE MARZANO CAUSAL TEACHER EVALUATION MODEL AND ITS IMPACT ON PROFESSIONAL PRACTICES

SUSAN KATHRYN GRAZIANO

Dissertation Committee
Barbara V. Strobert, Ed.D., Mentor
Anthony Colella, Ph.D.
Carolyn Sattin Bajaj, Ph.D.

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Education

Seton Hall University
2016
SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Susan K. Graziano, has successfully defended and made the required modifications to the text of the doctoral dissertation for the Ed.D. during this Fall Semester 2016

DISSERTATION COMMITTEE
(please sign and date beside your name)

Mentor:
Dr. Barbara Strobert

Committee Member:
Dr. Anthony Colella

Committee Member:
Dr. Carolyn Sattin-Bajaj

The mentor and any other committee members who wish to review revisions will sign and date this document only when revisions have been completed. Please return this form to the Office of Graduate Studies, where it will be placed in the candidate's file and submit a copy with your final dissertation to be bound as page number two.
Abstract

Teacher evaluation practices have been a common topic of discussion at the federal, state, and local levels in recent years. The literature asserts that teacher evaluation has a dual purpose: (1) to improve teacher instructional practice, and (2) to inform personnel decisions (retention and dismissal) (Donaldson & Papay, 2014; Gabriel, 2015; Garrett & Steinberg, 2015; Marzano & Toth, 2013). While districts are certainly utilizing evaluation instruments for the purpose of making personnel decisions, there is little research to support the claim that teacher growth is occurring as a direct result of teacher evaluation practices. It is essential that educational leaders consider teacher perception of teacher evaluation practices if the evaluation practices are to effectively address and fulfill a purpose of teacher evaluation: improvement of teacher instructional practice. Perception represents an individual’s reality and influences one’s interactions with and opinions of a particular phenomenon. For these reasons, the researcher explored teacher perceptions of a widely utilized evaluation model: The Marzano Causal Teacher Evaluation Model.

This study focused specifically on the ways in which teachers perceive the Marzano model has or has not influenced professional practices and relationships since its adoption in 2012 at a New Jersey high school. Further, this study is an investigation of teachers’ beliefs regarding the accuracy of their evaluations ascertained by the use of the Marzano Model.

The results of this study speak to the complexity of the change process, the importance of the perceived validity of the evaluation instrument, the influence of perceived credibility of and support from evaluators, and the overwhelming desire for
educators to hone and improve their craft. The results of this study also suggest that if administrators are viewed as coaches instead of solely as assessors, teacher evaluation practices have the potential of improving teacher practice and, ultimately, student achievement.
Acknowledgments

As I reflect upon the journey that has resulted in the pages that follow, I am filled with pride, but mostly appreciation for all who have helped me to achieve the goal that I set out to accomplish years ago. It is true that I have spent countless hours working toward the completion of the research that acts as the culmination of my academic experience at Seton Hall University. However, I would never have had the pleasure of this accomplishment without the unceasing support, guidance, dedication, and love from so many individuals.

To my dissertation advisor, Dr. Barbara V. Strobert, I thank her for inspiring me, challenging me, and supporting me throughout this process. Dr. Strobert has demonstrated such kindness and compassion, constantly reminding me why I am and will always be an educator. She will never know the impact that she has had on my life, both professionally and personally. There were many points when I experienced self-doubt, but Dr. Strobert seemed to know just when I needed those small words of encouragement. Throughout this process, she provided me with prompt feedback, probing questions, and invaluable insight. The pages that follow are a direct reflection of her support and guidance as my mentor. Dr. Strobert, you have nothing but my admiration and respect.

I would also like to recognize and express sincere appreciation for my dissertation committee, Dr. Anthony Colella and Dr. Carolyn Sattin-Bajaj, for their willingness to offer their time, support, and guidance throughout this process. Dr. Colella, you have been kind and gracious with your encouragement and assistance throughout the past year. Dr. Sattin-Bajaj, you instilled a love of qualitative research in me a few short months ago. Your passion in discovering human truths is contagious. You both have influenced these pages and have strengthened my belief in the value of educational research.
I also wish to thank the following four individuals who served on my panel of experts: Mrs. Stacey Oehler, Mr. Steven Ottinger, Mr. Brian Tonelli, and Dr. John Cafagna. Your insights helped me to craft a well-rounded research study. You have been gracious with your time and your thoughts, both of which were instrumental in my study.

I would like to express my sincere appreciation to all of my professors at Seton Hall University for imparting their wisdom and sharing their varied experiences in education. I have learned from each one of you. The leader I am today is a culmination of all of my professors along the way.

To the fourteen teachers who graciously shared their time with me: I value each and every one of you. I admire your openness and candor. I celebrate your professional efforts and wish to emulate your selflessness. I have gained so much from the short time I have shared with each of you. Our conversations have and will continue to shape the professional conversations that I have as an educational leader. This research study would not be possible without your willingness to participate.

I also wish to thank the administrators of the participating district. Your dedication to your profession is represented by your willingness to offer your school district as a research site for this study.

To Lauren Kerr, a person I consider a colleague and friend, I am indebted to you and will forever remember your kindness. You are the consummate professional and an all-around great person.

To Mrs. Donna Dickhart, a true friend and mentor, you have influenced my professional path more than you will ever know. If it weren’t for you, I would never have considered a career
in educational leadership. You saw something in me years ago that I, myself, could not see. I can only hope that one day I will touch someone else’s life the way that you have touched mine.

I wish to extend my love and appreciation to my parents, who have inspired me to set lofty goals and achieve them. You have taught me that the sweetest rewards are often preceded by the greatest challenges. I also want to thank my two sisters for their love, guidance, and support throughout the years. To my second family, the Grazianos, thank you for being an ear to listen, a shoulder to lean on, and a vote of confidence since I met you ten years ago. All of you have selflessly offered your time and support and have taught me what it means to be a family.

To my two beautiful daughters, Brooke and Hadley, I have felt the need to apologize to you since I began this journey. I have had to devote time and energy to classes, research, writing, revising, and editing. I know you have grown tired of asking the question: “Mommy, are you still writing your paper?” I have felt immense guilt every minute I have been away from you, but I hope that one day my experience will demonstrate the following: (1) My response to all of the goals that you put before you will always be: You can, and you will. (2) Never underestimate your own worth and abilities. (3) Your education is your very own rocket to the stars for which you are shooting; please don’t ever take it for granted. (4) I love you each taller than the tallest giraffe. I hope you are proud to call me your mom.

Finally, to my husband, Joseph, I don’t even know where to begin. No words could ever express my gratitude and admiration for all that you have done for me throughout this process. You have selflessly taken on any number of tasks to help alleviate my burdens. You have talked me off of the metaphorical “ledges” and provided words of encouragement when I needed them the most. You have never complained, nor have you questioned my efforts. You simply and completely supported me. You are a remarkable husband, father, confidant, companion, and
friend. You have taught me what it means to love without boundaries. I only wish that you
could see yourself through my eyes. My world is a better place because you are in it. I will
never take you or our happiness for granted. Thank you, thank you, thank you.
Dedication

I dedicate this work to all of the teachers who have decided to make it their life’s mission to awaken and inspire the genius in each one of their students. Please do not think that the endless hours you spend “off the clock” go unnoticed or unrecognized. Your words will live on long after you have left the classroom. May your legacy reflect positivity, generosity, and hope. May we all, as educators, continue to view our profession as an honor and a privilege, a unique opportunity to leave our footprints in the sand.
<table>
<thead>
<tr>
<th>Table of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chapter I. INTRODUCTION</strong> .................................................................................. 1</td>
</tr>
<tr>
<td>Context of the Problem ................................................................. 2</td>
</tr>
<tr>
<td>Statement of the Problem ............................................................. 4</td>
</tr>
<tr>
<td>Purpose of the Study ........................................................................... 6</td>
</tr>
<tr>
<td>Research Questions ........................................................................... 7</td>
</tr>
<tr>
<td>Theoretical Framework ..................................................................... 7</td>
</tr>
<tr>
<td>Design and Methodology ................................................................. 8</td>
</tr>
<tr>
<td>Significance of the Study ................................................................. 10</td>
</tr>
<tr>
<td>Limitations of the Study ................................................................. 12</td>
</tr>
<tr>
<td>Delimitations of the Study ............................................................... 13</td>
</tr>
<tr>
<td>Definition of Terms ........................................................................... 14</td>
</tr>
<tr>
<td>Summary .......................................................................................... 15</td>
</tr>
<tr>
<td><strong>Chapter II. REVIEW OF THE LITERATURE</strong> ......................................................... 17</td>
</tr>
<tr>
<td>Literature Search Procedures .......................................................... 17</td>
</tr>
<tr>
<td>Criteria for Inclusion and Exclusion of Literature ............................... 18</td>
</tr>
<tr>
<td>Historical Context of Instructional Supervision &amp; Teacher Evaluation ........................................... 19</td>
</tr>
<tr>
<td>Historical Context of Federal Legislation and Teacher Evaluation ................................................. 27</td>
</tr>
<tr>
<td>Theoretical Frameworks .................................................................... 30</td>
</tr>
<tr>
<td>Adult Learning Theory ...................................................................... 30</td>
</tr>
<tr>
<td>Social Cognitive and Social Cognitive Career Theory .............................. 32</td>
</tr>
<tr>
<td>Reflective Practice Theory ................................................................ 33</td>
</tr>
<tr>
<td>Standards-Based Evaluation in New Jersey ........................................... 35</td>
</tr>
<tr>
<td>Value-Added Measures and Student Growth Percentiles in Teacher Evaluation ........................................... 38</td>
</tr>
<tr>
<td>The Marzano Causal Teacher Evaluation Model ....................................... 41</td>
</tr>
<tr>
<td>Common Problems Associated with Teacher Evaluation ........................... 46</td>
</tr>
<tr>
<td>Effective Teacher Evaluation ............................................................ 50</td>
</tr>
<tr>
<td>Teacher Perception of Evaluation ..................................................... 52</td>
</tr>
<tr>
<td><strong>Chapter III. METHODS</strong> ............................................................................. 56</td>
</tr>
<tr>
<td>Background ........................................................................................ 56</td>
</tr>
<tr>
<td>Design ................................................................................................ 60</td>
</tr>
<tr>
<td>Profile of the Site .............................................................................. 61</td>
</tr>
<tr>
<td>Sampling ........................................................................................... 64</td>
</tr>
<tr>
<td>Data Collection .................................................................................. 66</td>
</tr>
<tr>
<td>Data Analysis .................................................................................... 71</td>
</tr>
<tr>
<td>Protection of Human Subjects and Ethical Considerations .......................... 72</td>
</tr>
<tr>
<td>Validity and Reliability .................................................................... 73</td>
</tr>
<tr>
<td>Role of the Researcher and Researcher Bias ........................................... 75</td>
</tr>
<tr>
<td>Summary ........................................................................................... 78</td>
</tr>
</tbody>
</table>
Chapter IV. RESULTS .........................................................................................79

Summary of Findings for Research Question 1 ...........................................80
Themes Research Question 1 .................................................................82
  Perceived Instructional and Pedagogical Emphases ..................82
  Approach to Lesson Planning .........................................................85
Summary of Findings for Research Question 2 ......................................87
Themes Research Question 2 .................................................................88
  Inter-rater Reliability .................................................................88
  Clarity of Expectations ..............................................................89
  Depth and Breadth of the Evaluation ..................................91
  Constructive and Useful Feedback .......................................93
  Perceptions of Proficiency Scales and Ratings ......................94
Summary of Findings for Research Question 3 .....................................98
Themes Research Question 3 .................................................................99
  Coaching Versus Evaluating ...................................................99
  Trust .........................................................................................101
  Understanding and Managing Change ................................104
Summary ..............................................................................................107

Chapter V. CONCLUSION AND RECOMMENDATIONS .............................108

Research Questions ..............................................................................109
Summary and Discussion of Major Findings ......................................110
  Findings Related to Research Question 1 .........................110
  Findings Related to Research Question 2 .........................113
  Findings Related to Research Question 3 .........................116
Findings Related to the Theoretical Framework ..............................118
Recommendations for Practice .........................................................122
Recommendations for Policy ...........................................................126
Recommendations for Future Research ..........................................128

REFERENCES .............................................................................................130

APPENDICES ..............................................................................................146

APPENDIX A. 2014 Marzano Teacher Evaluation Model Learning Map ....146
APPENDIX B: Letter of Solicitation ...........................................................148
APPENDIX C: Demographic Profile Questionnaire .........................151
APPENDIX D: Informed Consent Form ..................................................153
APPENDIX E: Certificate of Completion: Protecting
  Human Research Participants .........................................................158
APPENDIX F: Interview Questions ..........................................................160
APPENDIX G: School’s Permission Granted ........................................163
APPENDIX H: IRB Approval ................................................................165
APPENDIX I: Approval for Dissertation Proposal ..............................167
APPENDIX J: Lesson Plan Template ......................................................169
List of Tables

Table 1. Summary of Demographic Profile Information for Each Teacher ....................65
Table 2. Overview of the Procedures Used to Collect Data ..............................................67
Table 3. Sample Interview Questions Raised by Theoretical Framework ......................69
Table 4. Preliminary Set of Data-Driven Codes .................................................................72
CHAPTER 1
INTRODUCTION

Most schools in the United States require between 175 and 180 days of school, during which students receive between 900 and 1,000 hours of instruction (Hull, 2011). The quality of instruction, more specifically, the quality of the teacher planning and implementing instruction, has a direct impact on student achievement (Sanders, Wright, & Horn, 1997; Sanders & Horn, 1998; Darling-Hammond, 2000; Heck, 2009). For this reason, teacher evaluation practices have been a common topic of discussion at the federal, state, and local levels in recent years. Teacher evaluation has a dual purpose: (1) to improve teacher instructional practice, and (2) to inform personnel decisions (retention and dismissal) (Donaldson & Papay, 2014; Gabriel, 2015; Garrett & Steinberg, 2015; Marzano & Toth, 2013). Since 2008, a large majority of states across the United States have adopted new teacher evaluation practices and policies that include multiple measures of teacher quality, including professional practice, instructional practice, and student achievement and growth. In 2009, the National Education Association (NEA) published a study in which they identified 25 new or proposed laws and regulations regarding teacher evaluation (NEA, 2011). All 41 state applications for Race to the Top (RTTT) funding mentioned teacher evaluation (Learning Point Associates, 2010).

Twenty-eight states, including New Jersey, use teacher evaluation results to inform personnel decisions (Hull, 2014). If a school district wants to perform at its highest level, district administrators must make every effort to hire and retain high quality teachers and dismiss those who underperform, regardless of tenure status. The decision-making process is largely, if not entirely, influenced by teacher evaluation practices. Evaluation practices are intended to inform and, ultimately, improve teacher performance, but the question remains whether this is the case.
The current study explores teacher perception of Robert Marzano’s Causal Teacher Evaluation Model and the ways in which the evaluation model, adopted and modified by a school district in Southern New Jersey, affects and informs teachers’ instructional practices.

**Context of the Problem**

In 1983, *A Nation at Risk: The Imperative for Educational Reform*, a report published by the United States Department of Education, recommended that “Salary, promotion, tenure, and retention decisions should be tied to an effective evaluation system.” The report garnered widespread public attention in response to the claim that America’s schools were “mediocre,” falling behind competing nations that previously had paled in comparison. In the years that followed, the focus on accountability measures for student achievement and teacher performance was evident. In 2002, The No Child Left Behind Act (NCLB) called for nationwide accountability measures for student achievement through use of standardized assessments (U.S. Department of Education). In February of 2009, President Barack Obama signed the American Recovery of Reinvestment Act (ARRA) into law under which $4.35 billion was allotted for Race to the Top funding (U.S. Department of Education, 2009). The Race to the Top program called for a “comprehensive approach to education reform” that included implementation of a fair evaluation system that would “differentiate [teacher and principal] effectiveness using multiple rating categories that take into account data on student growth” (U.S. Department of Education, 2009, p. 9). Race to the Top, unlike its policy predecessors, recommended a direct link between teacher evaluation and student achievement.

In August of 2009, Weisberg et al. published a report titled *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness* in *The New Teacher Project* in which they identified a “fundamental crisis” that they coined the Widget
Effect, a phenomenon that “describes the tendency of school districts to assume classroom effectiveness is the same from teacher to teacher” (pp. 2-4). The researchers claimed that administrators viewed teachers as “interchangeable parts” that could easily be replaced by other professionals who would provide the same quality of instruction. The researchers also claimed that the root of the crisis is the fact that districts poorly implemented evaluation systems and failed to utilize information gleaned from the evaluations to make important personnel decisions, including the practice of retaining and supporting good teachers and dismissing teachers that are deemed ineffective. In an effort to reverse the Widget Effect, the researchers called for an evaluation system that clearly identified both effective and ineffective teachers through the incorporation of a fair and equitable evaluation system. The report further highlighted ineffective teacher evaluation practices, resulting in a heightened sense of immediacy in teacher evaluation reform.

After The Widget Effect was published, many states, including New Jersey, adopted new evaluation policies. On August 6, 2012, The Senate and General Assembly of New Jersey passed the TEACHNJ Act (P.L.2012, c. 26) that states the goal of the legislation is to “raise student achievement by improving instruction through the adoption of evaluations that provide specific feedback to educators, inform the provision of aligned professional development, and inform personnel decisions.” On March 6, 2013, AchieveNJ, the reported improved education evaluation and support system, was proposed to the State Board of Education (New Jersey Department of Education, 2013). The Board adopted the system on September 11, 2013. The legislation required school districts to adopt a state-approved evaluation system by the 2013-2014 school year. State guidelines required that the selected evaluation systems include Student Growth Objectives (SGOs), Student Growth Percentiles (SGPs), if applicable, and observation
data garnered from a trained observer to calculate teachers’ overall summative evaluation score, by which teachers were assigned a score that was intended to accurately reflect overall teacher quality.

A large body of educational research has been and continues to be devoted to student achievement. Schools are rated, ranked, recognized, and reproached as a result of published standardized measures of student achievement. Years of research have established a direct link between teacher quality and student achievement (Darling-Hammond, 2000, 2013, 2015; Garrett & Steinberg, 2015; Hanushek, 1992; Heck, 2009). Teacher effect is the most significant indicator of student achievement gains, both annually and cumulatively (Sanders & Rivers, 1996; Sanders, Wright, & Horn, 1997). While the ramifications of poor teacher quality are repeatedly studied and reported, the body of research on how evaluation systems affect teacher performance is a bit more indistinct. Should student achievement be directly tied to teacher evaluation? Should value-added measures (VAMs) be factored into teacher evaluation systems? These debates continue in the field of educational research, despite repeated attempts at the federal and state level to systematize teacher evaluation practices.

**Statement of the Problem**

While the most recent legislation regarding teacher evaluation was intended to improve teacher practice and, ultimately, student achievement through use of an evaluation system that employed multiple measures, it is still relatively unclear how the new evaluation systems adopted across the state have informed or improved teacher practice. Most districts in New Jersey chose one of the five most popular state evaluation models (numbers reflect a survey conducted in 2013): *Charlotte Danielson Framework for Teachers* (291 New Jersey school districts); *Stronge Teacher and Leader Effectiveness Performance System* (53 districts); *Mid-
Continent Research for Education and Learning (McREL) Teacher Evaluation Standards (45 districts); Marzano’s Causal Teacher Evaluation Model (44 districts); The Marshall Rubrics (32 districts) (Mooney, 2013). The remaining districts, including Newark Public Schools, chose to adapt the popular models in order to tailor the evaluation system to the perceived needs of the district.

While much attention has been paid to teacher evaluation systems, teacher tenure laws, and teachers unions throughout the past decade, few studies reflect the human component of the TEACH NJ Act and AchieveNJ. In 2015, Callahan and Sadeghi conducted a survey that asked respondents to indicate their perceptions of the value of teacher evaluation practices and their perception of the value that administrators place on the evaluation system (p. 46). Fifty-three percent of the respondents indicated that teacher evaluation practices did not influence or change their instructional practices, therefore negating the intended purpose of TEACH NJ. Researchers and school officials alike agree that current evaluation practices are not effective in improving teacher practice or student achievement outcomes (Popham & DeSander, 2014; Quinn, 2014). Why, in spite of all of the state and district-level changes in teacher evaluation practices, do many teachers perceive evaluation practices as minimally impactful?

Recent publications reflect a need for more research regarding teacher perception of evaluation practices. Current teacher evaluation systems were implemented across the state of New Jersey to both improve teacher practice and assess teacher quality. If teachers are not employing the suggestions of evaluators, the former is not likely to happen. Through this study, the researcher intends to explore whether the Marzano Causal Teacher Evaluation Model has influenced teacher practice since its adoption in a New Jersey high school in 2012. Further, this
study is an investigation of whether or not the Marzano Model has influenced change in teacher
cpyactice or teachers’ perception of their impact on student achievement.

**Purpose of the Study**

According to Holsinger (2010), the primary emphasis of supervision is to increase
student achievement through use of research-based instructional practices. Student achievement
is a primary component in the majority of current evaluation systems. Teacher effectiveness is
one of the most significant factors in student achievement (Garrett & Steinberg, 2015; Heck,
2009; Sanders et al., 1997; Stronge et al., 2007). The Marzano Model reflects this emphasis, as
41 out of the 60 (roughly 68%) of the elements in the Marzano Teacher Evaluation Model relate
to classroom strategies and behaviors. In the 2001 book *Classroom Instruction that Works:*
*Research-based Strategies for Increasing Student Achievement*, Robert Marzano identified the
following nine essential instructional strategies that have the greatest impact on student
achievement:

- Setting Objectives and Providing Feedback
- Reinforcing Effort and Providing Recognition
- Cooperative Learning
- Cues, Questions, and Advance Organizers
- Nonlinguistic Representations
- Summarizing and Notetaking
- Assigning Homework and Providing Practice
- Identifying Similarities and Differences
- Generating and Testing Hypotheses
Each of these strategies is reflected in Domain 1 of the Marzano Model (Classroom Strategies and Behaviors), in which all 41 aforementioned elements are listed. The composition of the model reflects the emphasis on teacher practice, but is teacher practice improved through use of the model? When addressing teacher quality, one critical difference between ineffective and effective teachers is the fact that effective teachers are reflective practitioners, continually seeking to develop and cultivate highly effective instructional practices (Black & Howard-Jones, 2000; Stronge, 2007). Do teachers perceive teacher evaluation as a meaningful tool for professional growth and development? The purpose of this study is to explore how teachers perceive the influence of the Marzano Causal Teacher Evaluation Model on their professional practices.

**Research Questions**

The following four questions were used to guide the phenomenological qualitative research design:

1. How, if at all, has the Marzano Causal Teacher Evaluation Model influenced and informed teachers’ instructional practices?
2. What are teachers’ perceptions of how accurately the Marzano Causal Teacher Evaluation Model reflects and captures their professional performance and capabilities?
3. What are teachers’ perceptions of the influence of the Marzano Causal Teacher Evaluation Model on the relationships with their administrators?

**Theoretical Framework**

The theoretical frames that guided my research were social cognitive theory, social cognitive career theory, adult learning theory, and reflective practice theory. Social cognitive
theory and social cognitive career theory address the fact that individuals are sentient, purposeful beings who partially develop and process individual understanding through social interactions. Social cognitive and social cognitive career theories guide my understanding of how social interactions contribute to the individual and organizational understanding of and reaction to a professional practice. Adult learning theory addresses the necessary components of an effective adult learning environment. Knowledge of this theory guides my understanding of how adults learn and how the presence or absence of workplace characteristics influences and impacts professional development. Reflective practice theory ties to both perception and adult learning. Reflective practice involves assessing one’s own perceptions and actions for the purpose of cultivating and growing one’s craftsmanship (Osterman, 1990). Reflective practice theory informs my approach to determine if and how the Marzano Causal Teacher Evaluation Model enhances teachers’ reflective practice.

**Design and Methodology**

After being granted permission from the superintendent and building principal, I interviewed 14 teachers from a large suburban high school in southern New Jersey. Teachers were selected through purposive sampling. All teachers in the school district were invited to participate in the research. The information was communicated through use of a simple questionnaire that asked teachers to identify their interest in participation, the grade level and content that they teach, and the number of years they have been teaching. In order to encourage participation, the questionnaire also included the fact that all of the informants, as well as the school district, would be assigned pseudonyms to protect the anonymity of the participants and the school community. Participants were selected from the volunteer pool that represented a
range of grade levels, content areas, and years of experience. The number of years of experience of each of the teachers varied from intermediate (4-9 years) to veteran (10 or more years).

A phenomenological qualitative research design was most suitable because I aimed to describe teachers’ experiences and perceptions of the Marzano Causal Teacher Evaluation Model. Phenomenological qualitative research focuses on the experiences and lived events of humans. Further, the phenomenological study seeks to synthesize commonalities of collective perceptions and experiences related to a particular phenomenon (Saldana, 201). In this case, the phenomenon being explored was teacher evaluation practices.

The research design included two methods of data collection: semi-structured interviews and document analysis. The semi-structured interviews were in person, face-to-face, approximately 35-45 minutes in length, and were recorded using an audio recording device. The interview protocol included scripted open-ended questions derived from a review of the literature and guiding theoretical frameworks coupled with follow-up questions when further probing was deemed necessary. After the research questions were developed, interview questions were created that addressed each of research questions. The inclusion of structured interview questions was necessary in this research study in order to explore the breadth and depth of the phenomenon of teacher evaluation practice. The incorporation of structured interview questions also increased the likelihood that the results of the interview would be generalizable. All interviews were then transcribed within one week of the interview.

In order to increase the validity and reliability of the interview protocol, a jury of experts was assembled to test the quality of the questions. The panel consisted of one professor experienced in qualitative design, a principal of a large suburban middle school that conducted
similar research, and two expert teachers. The members of this panel were not participants in the study.

I also completed a document analysis of public records and physical evidence. The public records that were studied included the following: board policies that are relevant and applicable to the teacher evaluation practices and protocols in the district in order to gain a more in-depth knowledge of the ways in which the instrument is utilized in this school district. Artifacts of professional development flyers, meeting agendas for department or faculty meetings, and training materials available to teachers addressing the Marzano Causal Teacher Evaluation Model were examined for the purpose of gaining a better understanding of how the school offered professional development to teachers, if at all.

Data analysis followed the following three linked subprocesses, as suggested by Miles and Huberman, (1984, 1994): (1) data reduction, (2) data display, and (3) conclusion drawing/verification. All three subprocesses occurred before data collection, during data collection, and after data collection. Data reduction involves the act of condensing the data through identification of codes and themes. Data display involves organizing and compressing data in order to permit conclusion drawing. Conclusion drawing is the culmination of the previous two processes during which the researcher draws and verifies conclusions drawn after a thorough analysis of the displayed data.

**Significance of the Study**

Currently, educational researchers have identified a distinct paradigm shift in teacher evaluation practices: from supervision to evaluation and from teacher behavior to student achievement (Marzano et al., 2011). The current literature supports the recommendation for a comprehensive teacher evaluation system that employs multiple measures to assess teacher
effectiveness and, ultimately, promote professional growth (Bigham & Reavis, 2001; Liu, 2010; Darling-Hammond, 2012). The literature also supports the fact that there is a direct link between teacher quality and student achievement. In fact, teacher quality is consistently identified as a factor that has a significant impact on student achievement (Garrett & Steinberg, 2015; Heck, 2009; Sanders et al., 1997; Stonge et al. 2007). In light of the recent changes in teacher evaluation practices in the state of New Jersey, it is important to understand if these evaluation practices have aided in increasing teacher quality and, therefore, student achievement. Few studies have explored teacher perception of the new evaluation systems, specifically the Marzano Causal Teacher Evaluation Model. Since the Marzano Causal Teacher Evaluation model is currently used in 44 school districts across the State of New Jersey and many others across the United States, it is essential to understand how this instrument is perceived by its primary audience: the teachers. Understanding how teachers experience the Marzano Causal Teacher Evaluation Model is fundamental in understanding how the instrument can be used effectively.

The current research aims to understand the beliefs of teachers regarding current evaluation practices in a district that has adopted the Marzano Model. While the New Jersey Department of Education (NJDOE) requires specific components of an evaluation system, districts still have the autonomy to decide how they are going to implement the selected evaluation models. The results of this study can inform and guide the practices of instructional leaders using the instrument to assess teacher effectiveness. In order to improve evaluation practices, we must first begin the conversation and actively listen to both the commendations and concerns of the teachers. If educational leaders have a better understanding of how teachers perceive the use of the instrument, they can begin to reflect upon their own practices in an effort
to increase teacher quality and make informed policy decisions that reflect consideration of the identified concerns.

**Limitations of the Study**

Limitations of the study were as follows:

1. I interviewed teachers from one suburban school district in southern New Jersey. The results of the study are limited to the perspectives of these teachers. These findings, therefore, may be specific to the participants in this school district rather than representative of a larger population.

2. While other New Jersey school districts adopted other state-approved models, this study addressed the perspectives of teachers currently being evaluated under the Marzano Causal Teacher Evaluation Model.

3. The size of the sample (14 total participants) is a limitation of the study. This sample size represents approximately 12% of the total staff at the research site.

4. The composition of the sample is an additional limitation. The sample consisted of two males and 14 females. For this study, the percentage of males in the total sample is roughly 14%. However, males comprise approximately 32% of the staff at the research site. Males represent 27% of the district teaching staff, while their female counterparts represent 73%. Comparable districts in the surrounding areas report slightly higher percentages of male faculty members.

5. Another limitation of the study is the fact that the researcher used interviewing as a data collection method. The researcher must make the assumption that the participants are entirely truthful in their responses.
6. In qualitative research, the role of the researcher is that of a human instrument of data collection (Denzin & Lincoln, 2003). As a human being, I bring with me my own experiences and perceptions of the research topic I am exploring. My personal biases and experiences are more likely to influence the results of a qualitative study.

7. Finally, I am supervisor in the district. While I do not directly supervise any of the faculty working in the building at which the current study was conducted, I work with the administrators that directly supervise the participants. This fact could influence the participants’ responses to the interview questions.

**Delimitations of the Study**

Delimitations narrow the focus of study and inform the reader of the parameters and boundaries that were deliberately considered by the researcher (Calabrese, 2009). The design of this study included the following as delimitations:

1. The study is limited to teacher perspectives of the Marzano Causal Teacher Evaluation Model and does not address other approved and adopted evaluation models in the state of New Jersey.

2. The study was narrowed to one suburban high school in southern New Jersey that enrolls students from Grades 9-12.

3. Teacher evaluation ratings were not considered for the current study. Teachers did, however, report their summative evaluation scores in the demographic profile questionnaire. One teacher was rated as Highly Effective, while the remaining 13 participants were rated as Effective for the 2015-2016 school year. The composition of the sample regarding summative evaluation ratings was representative of the teaching population in the district.
4. The study is limited to the perspectives of teachers without researching teachers’ impact on student learning.

**Definition of Terms**

The following definitions provide clarification for terms used throughout this study:

**Applying.** A term used on the Marzano evaluation rubric that indicates that a teacher is effectively using an instructional strategy and is monitoring for the desired effect in the majority of students. The number value associated with Applying on the rubric is 3.

**Beginning.** A term used on the Marzano evaluation rubric that indicates that a teacher is using a strategy during instruction but not doing so effectively. The number value associated with Beginning on the rubric is 1.

**Common Language of Instruction.** An evaluation system in which all educational professionals talk about instruction in the same way, thus creating a shared understanding of professional expectations of the teacher (Marzano, 2008).

**Design Questions (DQs).** Questions teachers ask themselves when they are designing instruction. The Marzano Causal Teacher Evaluation Model includes 9 Design Questions.

**Developing.** A term used on the Marzano evaluation rubric that indicates that a teacher is using an instructional strategy but not monitoring for the desired effect for the majority of students. The number value associated with Developing on the rubric is 2.

**Elements.** The instructional strategies that happen in the classroom. The Marzano Causal Teacher Evaluation Model incorporates 60 individual elements.

**High Expectancy Students.** Students that are expected to perform very well for any number of reasons.
Innovating. A term used on the Marzano evaluation rubric that indicates that a teacher is effectively using an instructional strategy and is adapting the strategy to meet the needs of all learners in the classroom. The number value associated with Applying on the rubric is 4.

Instructional Strategies. The techniques that teachers implement during classroom instruction.

iObservation. A web-based system that collects, manages and reports information gained from walk-throughs, teacher observations, and teacher evaluations.

Learners. Students.

Low-Expectancy Students. Students that are expected to perform below average for any number of reasons.

Not Using. A term used on the Marzano evaluation rubric that indicates that a teacher is not using a strategy that is called for during instruction.

Reflective Practice. The ability of teachers to assess their own practices and make necessary adjustments to improve instruction.

Summary

This chapter introduced the research study through the context and statement of the problem of teacher perception of evaluation practices. The purpose and significance of the study were also revealed. I identified the ways in which this study can contribute to the current body of literature as well as implications for education stakeholders, specifically teachers and educational leaders. I also provided the qualitative methodology that was utilized for this research study. Delimitations and limitations were discussed to reveal the deliberate exclusions of the study as well as the ways in which the methods of the study limit the results of the study. Finally, I provided a list of defined terms that are referenced throughout the study.
Chapter II provides a review of the existing literature on teacher evaluation, specifically addressing the historical context of teacher evaluation, current practices in teacher evaluation, evaluation policy in the state of New Jersey, and the effect of teacher perception of evaluation on instructional practices. Chapter III presents a detailed description of the research design of the study, with particular emphasis on the methodology, data collection, and data analysis process conducted for the study. Chapter IV presents the results of the data collected from the research design. Chapter V presents a discussion of the findings of the study, conclusions drawn from the study, and implications for future research.
CHAPTER II
REVIEW OF THE LITERATURE

Introduction

This chapter provides a review of existing literature related to teacher perception of the evaluation practices and their impact on reflective practice. The review begins with an overview and analysis of the history of teacher evaluation, examining trends and evolving theories in effective teacher evaluation practices. This is followed by an examination of legislation and policies relating to teacher evaluation practices. This section is followed by a review of relevant theories relating to teacher perception of evaluation practices, specifically adult learning theory, social cognitive theory, and reflective practice theory. Next, I present an overview of standards-based evaluation in the state of New Jersey to provide a more specific geographical and political context from which this research study was designed. This is followed by an overview and analysis of the Marzano Teacher Evaluation Model, with an emphasis on theoretical and research-based implications of the implementation of the model. The literature review concludes with a discussion of current issues in teacher evaluation practices, which guided me in developing specific research questions about the implementation of the evaluation model in the district studied.

Literature Search Procedures

I conducted a literature search for research and literature related to teacher perception of evaluation through accessing the following resources: Academic Search Complete, Proquest Education Journals, Educational Resource Information Center (ERIC) (EBSCO)(A&I), Proquest ERIC, Proquest Multiple Databases, Proquest Social Science Journals, Proquest Psychology Journals, Education Data Catalog (U.S. data.gov), Digest of Education Statistics (National
Center for Education Statistics), U.S. Department of Education data, Google Scholar, books published on teacher evaluation, books published on learning theory, the New Jersey Department of Education website, and the United States Department of Education website. The following search terms were used throughout the literature search process: teacher perception of evaluation, teacher evaluation, teacher evaluation in New Jersey, history of teacher evaluation, Marzano Causal Teacher Evaluation Model, teacher effectiveness, teacher evaluation policy, reflective practice theory, adult learning theory, organizational learning theory, social cognitive theory, and social cognitive career theory.

Most of the works consulted were peer-reviewed scholarly articles obtained from Seton Hall University’s library databases. Most works consulted were published within the last ten years; however, some works referenced were older due to the fact that they are primary sources that contributed to my knowledge of historical and theoretical considerations in teacher perception of evaluation.

**Criteria for Inclusion and Exclusion of Literature**

Studies were considered for inclusion for review, if the following criteria were met:

- English language literature and research articles published within the last 15 years, unless the work was historical or theoretical in nature
- Peer-reviewed journal articles
- Qualitative and quantitative scholarly research publications from peer-reviewed professional journals
- Evidence-based commentary in peer-reviewed journals
- Articles from respected education and educational research journals
- Books and book chapters on teacher evaluation
Books and book chapters on qualitative research
- Books and book chapters that approach the topic from a theoretical framework; conference papers
- Government reports on education
- Federal and state legislation as background and contextual information
- TEACHNJ and AchieveNJ resources from the New Jersey State Department of Education website

Studies were considered for exclusion from the review if the following criteria were met:
- Literature relating to the perceptions of higher education professionals regarding teacher evaluation
- Literature that is not written in English
- Research studies performed in non-public schools in the United States.

**Historical Context of Instructional Supervision and Teacher Evaluation**

When taking a holistic approach, one could examine the history of teacher evaluation in the United States beginning in the 1700s, a time period when clergy were solely responsible for evaluating, hiring, and firing teachers, who were considered servants to the community.

Shinkfield and Stufflebeam (2007) identified five major periods in teacher evaluation: (1) the Pre-Tylerian Period (before 1930), (2) the Tylerian Age (1930-1945), (3) the Age of Innocence (1946-1957), (4) the Age of Realism (1958-1972), and (5) the Age of Professionalism (1973 – present day). In order to understand the current state of teacher evaluation, it is best to begin with the latter part of the 19th century and early part of the 20th century, the Period of Scientific Management, an era that is, in part, defined by the conflicting views of John Dewey and Edward
L. Thorndike (Marzano, Frontier, & Livingston, 2011). The two conflicting theories of Thorndike and Dewey are most representative of the Pre-Tylerian Period.

In her article “The Plural Worlds of Educational Research,” Ellen Condliffe Lagemann (1989) stated: “One cannot understand the history of education in the United States during the twentieth century unless one realizes that Edward L. Thorndike won and John Dewey lost” (p. 185). Lagemann herself admitted that she made this statement “only in part to be perverse,” indicating that there was some truth in the statement, however simple it may seem. While Dewey emphasized the importance of fostering citizenship and democratic ideals in the classroom, Thorndike valued measurement as the more scientific and therefore more reliable approach to schooling (Marzano et al., 2011). Thorndike’s theories, inspired by Frederick Taylor (1911), emphasize the importance of efficiency and precision within the structure of the school.

Dewey and Thorndike also had differing views of teacher evaluation. Dewey believed that in order to foster collaboration and true professional growth, educational organizations should avoid “rigid, institutionally bound roles” (Lagemann, 1989, p. 201). The only difference between the supervisor and the teacher is the primary role each fulfills. Thorndike’s philosophies, however, are a stark contrast to Dewey’s collaborative model. In his book Public School Administration, Ellwood Cubberley (1916) applied Frederick Taylor’s principles of scientific management and Thorndike’s theories of effective educational programs through use of an analogy that identified schools as factories and children as products:

Our schools are, in a sense, factories in which raw products (children) are to be shaped and fashioned into products to meet the various demands of life. The specifications for manufacturing come from the demands of the twentieth-century civilization, and it is the
This statement is representative of a scientific approach to education: each professional (teachers, principals, superintendents, etc.) has an assigned role through which he is expected to produce an expected or desired outcome. Further, these outcomes should be measured through use of data compilation and analysis (Cubberley, 1916; Marzano et al., 2011).

The work of Thorndike and Cubberley continued to influence public education throughout the Great Depression. William Wetzel (1929) proposed use of student achievement measures to determine the effectiveness of the school, a practice that is not unfamiliar in today’s educational landscape (Marzano et al., 2011). Wetzel further asserted that the assessments by which student achievement is measured must bear evidence of reliable student ratings. According to Wetzel, the effective instructional supervisor utilized reliable student data to make reasonable and accurate recommendations and commendations of teacher practice. The onus, however, was placed primarily on the student during this time period, and an emphasis on teacher evaluation was lacking from the literature of this time period (Shinkfield & Stufflebeam, 1995).

Cubberley and Wetzel’s scientific approach to educational evaluation maintained relative popularity until the Post-World War II era. However, the Tylerian Age, as identified by Shinkfield and Stufflebeam (2007), influenced evaluation policy during this time period as well. Ralph W. Tyler, the individual from whom the Tylerian Period was established, emphasized the importance of instructional objectives and student outcomes. Tyler encouraged teachers to identify the skills and strategies they intended students to achieve or master. Evaluators could then assess teacher effectiveness by determining how well the students achieved the intended
objective (Nowakowski, 1983). This approach to instructional evaluation influenced Tyler’s predecessors and is utilized in some capacity in each of the New Jersey state-approved evaluation models today.

During the period following World War II, there was a distinct shift in evaluation practices. Classified as the Age of Innocence by Shinkfield and Stufflebeam (2007), this era in educational evaluation was characterized by the overwhelming sense of relief following a victory in World War II. The American economy was on the mend and the public education system was flourishing with increased academic course offerings, supported by increased federal funding (Madaus, Scriven, & Stufflebeam, 1984). The scientific approach to education was no longer widely accepted. Rather, the literature began to focus on the teacher as an individual and the instructional leader as one who would tend to both the professional and emotional needs of the teacher (Marzano et al., 2011). In an article published in Educational Leadership, Elsie Coleman (1946) described the modern approach to “supervisory visits”:

So the supervisor visits the teacher in many ways, within and without the school understanding the needs and possibilities, building human relationships, using procedures that are cooperatively planned and evaluated, becoming counselor, guide, friend. (167)

Ethel Thompson (1952) recommended a collaborative approach to supervision through which the supervisor is dependent upon the teacher to determine the applicability and effectiveness of both old and new instructional strategies. This multifaceted approach to supervision was a far cry from the scientific management approach that was preferred in years past. Coleman (1946) and Thompson (1952) presented an approach to supervision that was collaborative and humanistic, with an emphasis on the social-emotional component of evaluation.
The Post-World War II era in evaluation was also characterized by an increase in supervisory responsibilities, resulting in a broadened job description of the supervisor (Marzano et al., 2011). During this era in education, practitioners and researchers alike agreed that the instructional leader should approach classroom visitations, teacher evaluation, and interactions with teachers with the mentality that supervisors and teachers were a team working toward a common goal: student success and achievement. Specific processes through which the evaluator and teacher should go about achieving this goal remained ambiguous, however, until the era of clinical supervision.

1957 marked another distinctive shift in evaluation. In 1957, the Russians successfully launched Sputnik I, sparking a nationwide educational crisis (Hogan, 2007). In 1958, Congress enacted the National Defense Education Act (NDEA), through which millions of dollars were invested in a structural and programmatic overhaul of American public education. New curricula were developed, and evaluations were utilized to determine the success of the new curricula (Hogan, 2007). The era marks the Age of Realism, as a compulsory evaluation of American public education ensued as a direct result of the fear that Americans were falling behind other developed countries.

During this time period, clinical supervision was developed and quickly became a popular model in the field of education (Marzano, et al., 2011). Morris Cogan, a professor at Harvard’s Master of Arts in Teaching program, developed clinical supervision in the 1950’s with a group of colleagues working at the Harvard M.A.T. program. The model was developed with the intention of providing a systematic approach to instructional supervision. The model was made popular by Robert Goldhammer’s book titled Clinical Supervision: Special Methods for the Supervision of Teachers (1969), in which Goldhammer developed a five-phase process of
supervision: Phase 1 – Preobservation Conference; Phase 2 – Classroom Observation; Phase 3 – Analysis (of data obtained from the observation); Phase 4 – Supervision Conference; Phase 5 – Analysis of the Analysis (analysis of the performance of the supervisor) (Goldhammer, 1969).

Clinical supervision, as envisioned by Goldhammer, provided educational professionals, specifically instructional leaders, with a clearly-defined process through which supervisors could (in theory) help teachers foster more effective instructional practices. However, the process itself became a formulaic approach to classroom supervision, lacking the enriched professional dialogue accompanying the process, a component that was necessary for achieving the intended purpose of its implementation (Marzano et al., 2011).

These five steps are reminiscent of many current supervisory practices, which employ most, if not all, of the five phases identified in clinical supervision. According to Reavis (1976), clinical supervision “rests on the conviction that instruction can only be improved by direct feedback to a teacher on the aspects of his or her teaching that are of concern to that teacher (rather than items on an evaluation form or items that are pet concerns of the supervisor only)” (p. 360). Targeted feedback is a component of many approved current teacher evaluation models. Unfortunately, the model drove the practice instead of the reverse.

The final identified era in evaluation was the Age of Professionalism (Shinkfield & Stufflebeam, 2007). This era was characterized by a storm of professional publications and journals including *Educational Evaluation and Policy Analysis, Studies in Educational Evaluation, and Evaluation Review* (Hogan, 2007; Stufflebeam, et al., 2000). Madeline Hunter, an esteemed education practitioner, influenced the practices of teachers and supervisors alike throughout this time period through her behaviorist approach to effective classroom instruction. Hunter is most famous for her seven-step lesson design model that includes each of the
following: anticipatory set, objective and purpose, input, modeling, checking for understanding, guided practice, and independent practice (Hunter, 1984). Hunter also contributed to current practices in instructional supervision. Hunter viewed principals as instructional “coaches,” individuals who have the knowledge and expertise to strategically and deliberately improve teacher instruction (Brandt, 1985). In her article “Six Types of Supervisory Conferences,” Hunter (1980) identified two discrete functions of supervisory conferences: (1) the conference must “promote growth in effective instruction” and (2) the conference serves as an evaluation of the teacher (p. 408). Hunter argued, however, that evaluative conferences should be the culmination of several supervisory visits, through which teachers and supervisors engage in a preconference before the observation and a post-conference following the observation.

Hunter (1980) also argued that teacher evaluation characterized a process through which teachers were placed on a continuum from “unsatisfactory” to “outstanding” and provided teachers an opportunity to reflect upon the summative evaluation rating through examination of multiple data points. Hunter’s model dominated the landscape of teacher evaluation during the 1980s and popularized the trend for instructionally focused staff development (Brandt, 1995). Clinical supervision provided the framework from which supervisors should structure the process of teacher evaluation, while Hunter’s 7-Step Model informed conversations between teachers and administrators about instruction.

In June 1984, the RAND group published a report on teacher evaluation practices titled *Teacher Evaluation: A Study of Effective Practices*, a study that surveyed 32 school districts in an effort to uncover effective teacher evaluation practices (Wise et al., 1984). Ultimately, the researchers chose to complete a case study of evaluation practices in four distinctly different school districts: Salt Lake City, Utah; Lake Washington, Washington; Greenwich, Connecticut;
and Toledo, Ohio. After studying the four districts, Wise et al. (1984) identified four characteristics of effective evaluation implementation: organizational commitment, evaluator competence, teacher-administrator collaboration, and strategic compatibility. In summary, a school organization must be well versed in and committed to an evaluation system in order to utilize the system to its fullest capacity. Likewise, a district must select and/or develop an evaluation system that best aligns with organizational and community goals. In doing so, the district is more likely to implement an evaluation system equitably and efficiently.

After President Ronald Reagan’s National Commission on Excellence in Education report titled *A Nation at Risk: The Imperative for Educational Reform* was published in 1983, the American public became increasingly aware of and interested in teacher evaluation practices. The report cited an over 50-point decrease in the verbal section and nearly 40 points in the mathematics section of the Scholastic Aptitude Test (SAT) score. The committee recommended that teachers demonstrate competence in academic discipline as well as recommending performance-based pay for teachers, highlighting the need for skilled professionals in the teaching profession. Teacher accountability became a new “buzzword” in the public education sector, forcing school districts to reevaluate teacher evaluation.

In 1986, the Carnegie Forum on Education and Economy released the report titled *A Nation Prepared: Teachers for the 21st Century: The Report of the Task Force on Teaching as a Profession*, in response to the 1983 publication of *A Nation at Risk*. The report called for more rigorous standards in the education profession through the creation and incorporation of The National Board for Professional Teaching Standards. The report also indicated a need for the mobilization of resources to redefine the teaching profession through restructuring teacher preparation programs at the university level and revising teacher pay scales to be more
competitive with other professions. *A Nation Prepared* introduced the concept that teachers should be held accountable for student progress and performance, a concept that is reflected in most current evaluation systems (*A Nation Prepared: Teachers for the 21st Century*, 1986). *A Nation at Risk*, the RAND study, and *A Nation Prepared* were driving forces behind the overhaul of teacher evaluation practices in the United States. The following section examines the historical context of federal laws and initiatives that have propelled the high-stakes teacher evaluation policies across the United States.

**Historical Context of Federal Legislation and Teacher Evaluation**

Prior to the 1965 adoption of President Lyndon B. Johnson’s Elementary and Secondary Education Act (ESEA), the federal government had a peripheral role in public education (Cusick, 2014; Jennings, 2015). The primary focus of the law was educating America’s disadvantaged youth. Specific funds (Title I and Title II) were allocated to each school district in an effort to provide comparable opportunities for the economically disadvantaged students. The law, which has been regularly reviewed and renewed over the past five decades, laid the foundation for other federal laws that sought educational equity, including the Individuals with Disabilities Education Act (IDEA) of 1975. While this policy called for a regular review of curriculum and allocation of educational resources more so than a rigorous teacher evaluation system, the act was the first of its kind with regard to federal involvement in education policy (Jennings, 2015).

Public attention on student achievement and the quality of the American public education system continued at a steady pace until President George W. Bush signed the No Child Left Behind Act (NCLB) into law on January 8, 2002. The act was an update to ESEA and increased the federal government’s role in public education through a series of requirements that districts must meet in order to continue to receive Title I funding (Klein, 2015). In an attempt to enhance
all students’ academic performance, NCLB mandated that states develop a test-based student assessment program and publish data collected from the assessments. In terms of teacher quality, the Act required all teachers to earn “highly qualified” status by the 2005-2006 school year. In order for teachers to earn “highly qualified” status, they must have the following: (1) a bachelor’s degree, (2) full state certification or licensure, and (3) prove that they know each subject they teach (U.S. Department of Education, 2004). Under NCLB school-wide student achievement accountability was emphasized, as schools were required to make Adequate Yearly Progress (AYP); failure to do so for two consecutive years led to increased scrutiny under the state education agency and required underperforming districts to offer waivers by which students could enroll in a higher performing school (Hamilton, 2007; Klein, 2015). NCLB propelled the current era in education of high-stakes testing as a measure of school accountability.

In 2009, Congress authorized $4.35 billion in funding under the American Reinvestment and Recovery Act (ARRA) for a grant program called the Race to the Top (RTTT) Initiative, the largest competitive grant program ever instituted by the federal government (Herlihy, et al., 2014; U.S. Department of Education, 2009). The RTTT program cited four core education reform areas:

- Adopting standards and assessments that provide students the foundation to be successful in postsecondary institutions as well as the workforce
- Building data systems that measures student growth and inform teachers and administrators about how they can improve instruction
- Recruiting, developing, rewarding, and retaining highly effective educators, especially in high-needs districts
• Turning around the lowest-achieving schools (U.S. Department of Education, 2009).

RTTT emphasized again the importance of improving America’s public schools, yet this time partial emphasis was placed on teacher quality, a component that was not previously highlighted in federal education policy.

After it became clear that few, if any, states were on target to meet the rigorous requirements of NCLB, the ESEA Flexibility Program of 2011 awarded waivers (Popham, 2013). Forty-three of the 45 states that applied for the waivers were approved (U.S. Department of Education, 2016). In order to be approved, districts had to prove that they were successfully employing strategies to work towards the four core education reform areas cited in RTTT. The implementation of these waivers further solidified districts’ efforts to implement a comprehensive teacher evaluation system that included measures of student growth.

Most recently, President Barack Obama signed the Every Student Succeeds Act on December 10, 2015, which served as a reauthorization of the ESEA of 1965 and was described by the United States Department of Education as “the bipartisan bill to fix No Child Left Behind” (U.S. Department of Education, 2016). Of the 10 listed priorities, the following bear importance on teacher evaluation policies:

• Annual statewide assessments of all student learning
• Student performance targets and school ratings
• Accountability, interventions, and supports for struggling schools
• Competitive program to evaluate and reward effective educators (based on student learning) in high-needs schools
As with its policy predecessors, student achievement is the core priority of ESSA. The act maintains the emphasis, however, on teacher accountability and continues the conversation, or what many would call the debate, about merit-based pay. Past and present federal bills tell the story of the federal government’s increasing involvement in public education. More recent history suggests that high-stakes teacher evaluation is an ever-growing trend that appears not to be diminishing.

**Theoretical Frameworks**

**Adult Learning Theory**

In order to fully understand how teachers perceive the influence of teacher evaluation practices, one must first examine Adult Learning Theory. Simply put, adults learn differently than do children due to the fact that adults are often “relearning” the content being presented. Adults are often required to reconcile previous beliefs with newly constructed beliefs. Mezirow (1981) explored the work of Jurgen Habermas, from which he coined the term *transformative learning*. Habermas (1971) identified three generic domains of human interest: work knowledge (the ways in which one controls and manipulates his own environment), practical knowledge (social interaction) and emancipatory knowledge (self-knowledge). Work knowledge, also referred to as “technical knowledge,” is the lowest form of learning. Examples of work knowledge include rote memorization and knowledge of rules and expectations. Teachers may have knowledge of their district’s evaluation policies and practices, for example, but the focus remains on the how and not the why. Practical knowledge involves understanding of social norms. Teachers may understand and exhibit behaviors that they feel are appropriate when interacting with a student. This type of knowledge is developed through knowledge of the organization’s cultural and social norms. Emancipatory knowledge fosters a deep understanding
of the information presented. Teachers would know and understand why the evaluation policies are being implemented and how these policies impact them on the individual and collective levels (Kitchenman, 2008; Prayer, 1993). Mezirow (1981) eventually revised Haberman’s three types of learning: work became instrumental; practical became dialogic; and emancipatory became self-reflective (Kitchenman, 2008).

Mezirow (1981) built upon the work of Habermas, coining the term perspective transformation. Mezirow (1981) argued that perspective transformation was central to Habermas’s third learning domain. Mezirow (1981) identified the following ten elements of perspective transformation:

1. A disorienting dilemma
2. Self-examination
3. A critical assessment of personally internalized role assumptions and a sense of alienation from traditional social expectations
4. Recognizing that one’s problem or dilemma is shared and is not exclusive or private to the individual
5. Exploring options for new ways of acting
6. Building competence and self-confidence in new roles
7. Planning a course of action
8. Acquiring knowledge and skills for implementing one’s plan
9. Provisional efforts to try new roles and assess feedback
10. Reintegration into society on the basis of conditions established by the new perspective
Recent legislation in teacher evaluation practices in the State of New Jersey (discussed in detail in the next section) is distinctly different from previous evaluation practices. In order for teacher evaluation practices to be successful, one must first acknowledge the fact that this legislation dictated a paradigm shift in the field of education and then recognize the process through which individuals reconcile past practices and beliefs with current practices and beliefs.

**Social Cognitive Theory and Social Cognitive Career Theory**

In a qualitative study, the central consideration is perception, how individuals experience or “see” the world in which they live. According to *The Sage Encyclopedia of Qualitative Research Methods*, “Perception, which is mediated through the interconnectedness of the mind and body, is an individual’s access to experience and interpretation in the world” (p. 606). In order to understand how teachers experience and perceive the Marzano Evaluation Model, one must first understand how individuals learn in a collaborative professional setting. Social Cognitive Theory (Bandura, 1977; 1986) indicates that individuals learn from one another. Human beings are “sentient, purposive beings” whose actions are the result of the combined influence of personal factors, environmental factors, and behavioral factors (Bandura, 2001).

Self-efficacy, for example, refers to an individual’s beliefs in his or her capabilities of performing a particular function or task (Bandura, 1977, 1986, 1997). Self-efficacy is a personal factor that contributes to an individual’s perception of the professional environment and his or her perceived ability to grow within the organization to meet the proscribed professional standards. Self-efficacy is not simply a singular static trait, but rather a dynamic and ever-evolving belief that relates to one’s self-view of personal capabilities in performing a particular task, such as designing and implementing an effective lesson or attempting a new pedagogical technique (Lent & Brown, 1996). In order for an evaluation system to inspire true growth in the
teachers, it is essential to understand and implement strategies to increase self-efficacy within the school building. While Bandura also discussed the impact of outcome expectations, the anticipated outcome as the antecedent of a specific behavior, he believed that self-efficacy was a more influential behavior determinant (Lent & Brown, 1996). For example, teachers may associate a positive outcome with earning a master’s degree, but if they feel as though they are not well equipped to handle the demands of an advanced degree, they will likely not attempt to do so.

Social cognitive career theory, a derivative of social cognitive theory, posits that there are two primary aspects of career development: (1) the level of attainment individuals achieve in their professional tasks and (2) the degree to which they persist at a particular work activity (Lent & Brown, 1996). People are active agents in their career development, and behavior associated with career development is often flexible and, more importantly, affected by change efforts (Brown & Brooks, 1984; Lent & Brown, 1996). Recent litigation indicates that teacher evaluation systems are designed to act as change efforts in the field of education, promoting professional growth while ensuring professional accountability. While self-efficacy is not a substitute for ability (hence, the need for professional accountability), social cognitive theory, as well as social cognitive career theory, provides those tasked with creating and implementing teacher evaluation processes and policies invaluable insight into how teacher evaluation can positively impact professional development and teacher effectiveness.

**Reflective Practice Theory**

An essential component of adult learning is reflection (Boud, Keogh, & Walker, 1985; Dewey, 1938; Mezirow, 1981). Reflection relates to an individual’s cognitive processes as he or she becomes conscious of, understands, analyzes, and critiques assumptions, beliefs, or emotions.
In order to fully understand reflective practice theory, one must first understand organizational learning theory (OL), the theoretical perspective from which reflective practice theory was developed. An organization is defined as a “relatively long-lasting system of individuals and tasks that pursues specific goals” (Kluge & Schilling, 2003). Schools are, by definition, organizations. The individuals are the education professionals (teachers, administrators, staff members), and one common goal of any school building or district (as paraphrased from countless building and district vision statements) is to produce capable citizens that are college and career ready. OL is based on the premise that organizational learning follows the pattern of the establishment of the following four “I’s”:

- Intuiting
- Interpreting
- Integrating
- Institutionalizing (Hilden & Tikkamakki, 2013)

Each of the four I’s, however, cannot exist independently, nor can they exist without the incorporation of reflective practice.

Several theorists have contributed to reflective practice theory. The first, and, arguably, the most well-known theorist to study reflective practice was John Dewey. As discussed previously, Dewey believed that individuals constructed knowledge from experience. Dewey’s perspective asserts that through immersing oneself into professional experiences, the practitioner is able to “chunk” the learning experience in preparation for reflective practice (Shulman, 1998).

Borton (1950) posed a series of three questions to ask any practitioner: What? So What? Now What? Essentially, these questions inspired reflective thought through consideration of how the experience could be improved when encountered again.

Kolb (1976) introduced experiential learning theory, through which he depicted a four-stage cycle of learning, consisting of concrete experience, observations and reflections, formation of abstract concepts and generalizations, and testing implications of concepts in new situations. Essentially, in order to learn new processes or concepts, one must have the ability to
experience a new situation without bias, reflect upon this experience, connect the experience to abstract concepts and general understandings (theories), and use the new theories to solve real-world problems (Kolb, 1976; Kolb & Frey, 1979).

Adult learning is distinctly different from youth learning in the sense that adults are often reconstructing beliefs and ideas as opposed to developing new knowledge. Consequently, reflective practice is an essential component of adult learning and is essential to understand when examining the potential for or existence of professional growth.

**Standards-Based Evaluation in New Jersey**

Standards-based teacher evaluation is an evaluation process through which teacher performance is assessed and measured against a proscribed set of professional standards (Heneman et al., 2006). Standards-based evaluation systems provide an enriched picture of teacher performance through multiple measures. The recommended measures include multiple classroom observations, artifacts of student and teacher samples submitted by the evaluator and the teacher, and archived lesson plans submitted by the teacher (Milanowski, Kimball, & White, 2004; Danielson, 1996).

The New Jersey Department of Education (NJDOE) lists 11 total professional standards for teaching that are further categorized into four domains: Domain 1 – Learning and the Learner (Standards 1-3); Domain 2 – Content Knowledge (Standards 4-5); Domain 3 – Instructional Practice (Standards 6-8); and Domain 4 – Professional Learning (Standards 9-11). The New Jersey State Department of Education provides specific performances (actions of the teacher), essential knowledge (understandings of the teacher), and critical dispositions (behaviors of the teacher) associated with each of the 11 standards.
For example, Standard One is labeled “Learner Development” and is a component of Domain 1, “The Learner and Learning.” One expected performance of this standard is identified in the third professional expectation: “The teacher collaborates with families, communities, colleagues, and other professionals to promote learner growth and development.” In order to meet this expectation, the following is identified as essential knowledge: “The teacher understands the role and impact of language and culture in learning and knows how to modify instruction to make language comprehensible and instruction relevant, accessible, and challenging.” Finally, a critical disposition that demonstrates knowledge and understanding of this standard is identified as the following: “The teacher values the input and contributions of families, colleagues, and other professionals in understanding and supporting each learner’s development” (New Jersey Professional Standards for Teachers Alignment with InTASC, 2014).

The policy aligns with The Interstate Teacher Assessment and Support Consortium (InTASC), a consortium of state education agencies and national education organizations designed to develop high-quality professional standards for the teaching profession (New Jersey Professional Standards for Teachers Alignment with InTASC, 2014).

As the focus on teacher quality began to infiltrate the New Jersey policy agendas in recent years, teacher evaluation models began to gain popularity. Charlotte Danielson’s Framework for Teaching, Mid-Continent Research for Education and Learning (McREL) Teacher Evaluation Standards, Stronge’s Teacher and Leader Effectiveness Performance System, and Marzano’s Causal Teacher Evaluation Model are the most popular state-approved evaluation models in the State of New Jersey (New Jersey Department of Education). These models are state-approved and aligned with the professional teaching standards identified by the New Jersey Department of Education. While the recommended implementation of each evaluation model is
different, the intention remains the same: to improve instruction and student achievement, the primary purpose of teacher evaluation (Danielson, 1996; Danielson & McGreal, 2000; Heneman, et al., 2006; Marzano, 2011).

The TEACHNJ Act, adopted on August 6, 2012, was “rolled out” gradually through two rounds of regulations. The first round was intended to help districts prepare to implement improved evaluation systems in the 2013-2014 school year; the second round of regulations was intended to help districts implement the approved evaluation system in the 2013-2014 school year, with a projected adoption date of November 2013.

The state-approved policy required districts to move from a binary system of evaluation through which teachers were either performing or not performing a particular pedagogical skill to a system of multiple measures that included teacher practice (primarily through classroom observations), progress toward a Student Growth Objective (SGO) created by the teacher and approved by the evaluator, and in the cases of tested subject areas (Language Arts Grades 4-8 and Mathematics Grades 4-7), a median Student Growth Percentile (mSGP). During the year that the this study was conducted, teachers that taught non-tested areas were evaluated by a combined score that was comprised of 80% teacher practice (classroom observations) and 20% SGO. Teachers that were assigned an mSGP were evaluated by a combined score that was comprised of 70% teacher practice, 20% SGO and 10% mSGP. The weights of the percentages changed again for the 2016-2017 school year. Non-tested teachers are now evaluated by a combined score of 85% teacher practice and 15% SGO. Teachers assigned an mSGP are evaluated by the following weights: 55% teacher practice, 30% mSGP, 15% SGO.

A summative rating based on the aforementioned weights is calculated for each teacher. Each of the scores is calculated from a scaled rubric of 1-4 (1 = Ineffective; 2 = Partially
Effective; 3 = Effective; 4 = Highly Effective). Teachers who score below a 2.65 are assigned a Corrective Action Plan (CAP) through which the evaluator/s establish specific goals for growth and improvement. Teachers placed on a CAP for two consecutive years are eligible to be brought up on tenure charges (AchieveNJ, 2014).

**Value-added Measures and Student Growth Percentiles in Teacher Evaluation**

New Jersey’s teacher evaluation policy requires multiple measures of teacher effectiveness, one of which is similar to a value-added measure (VAM). Value-added measures, or growth measures, are incorporated to determine a student’s growth from one year to the next and, in turn, measure a teacher’s effectiveness. Barack Obama’s Race to the Top initiative required districts to “ramp up” teacher evaluation systems; as a result, many districts incorporated state standardized test scores as part of the teacher evaluation system. The quantitative value of the VAM is most often calculated by taking the student’s standardized test score, accounting for distinguishing characteristics of the population (minorities, students who are economically disadvantaged, etc.), and determining an approximate anticipated score for the next school year (Garrett, 2011). The validity and reliability of VAMs is hotly debated between teachers’ unions and policymakers. Many out-of-school factors are not and cannot be accounted for when calculating the expected growth of a student (homelessness, hunger, learning disabilities), thereby limiting the accuracy of the calculation (Garrett, 2011).

The New Jersey Department of Education, however, does not incorporate VAMs, but rather SGPs, into teacher summative ratings. SGPs differ from VAMs in the sense that SGPs are calculated by taking the median score of student standardized test scores, while the formula used to calculate VAMs is determined by the comparison between a student’s anticipated and earned score, based on the student’s previous score and the demographic data. Researchers have argued
that VAMs are more accurate than SGPs in determining teacher effectiveness (Walsh & Isenberg, 2015). In an effort to increase validity and reliability of the SGPs, the New Jersey Department of Education created peer school comparison groups. The New Jersey Department of Education defines peer schools as “schools that have similar grade configurations and are educating (or held accountable for) students with similar demographic characteristics” (NJDOE, 2015). Each peer school comparison group consists of approximately 30 schools with similar populations of the following:

- Percentage of students that are economically disadvantaged (derived from percentage of students that are eligible for free or reduced-price lunch)
- Percentage of students that are limited English Language proficient
- Percentage of students that are in special education
- Grade span of the school (elementary, middle, high or vocational high school)

(NJDOE, 2015).

Regardless of the existence of peer groups and subgroups, an SGP is a measure that is a single data point and not free of variables that cannot be controlled for, including those dictated by personal and environmental factors (David, 2010; Koretz, 2008; Schneider, 2012). The State of New Jersey continues to utilize information gained from mSGPs (median student growth percentiles) to calculate summative ratings for teachers (David, 2010). As previously noted, during the year that this study was conducted, 10% of a teacher’s summative evaluation score was derived from the teacher’s mSGP. The mSGP score is now weighted as 30% of a teacher’s overall summative evaluation score, reflecting a substantial increase in the emphasis on student assessment data. Interestingly, in the first year of implementation of mSGPs (2013-2014 school year), 30% of a teacher’s summative evaluation score was based on the mSGP; the New Jersey
Department of Education adjusted the calculation for the short time in between. The mSGP is certainly a consideration for New Jersey teachers, as the score directly influences the teachers’ evaluation, and potentially the teachers’ perceptions of the accuracy and value of the evaluation system.

Another concern that critics of mSGPs have raised is the fact that roughly half of the nation’s teachers are not assigned an SGP because they currently teach untested areas (Popham, 2013). In order for teachers to be assigned an mSGP score, they must: (1) be assigned to a fourth to eighth grade Language Arts or Math course for at least 60% of the school year prior to the administration of the standardized test, and (2) be assigned 20 unique students that have been enrolled for at least 70% of the course prior to the test administration (NJDOE, 2015). Special area teachers (physical education, art, music, technology, foreign language, etc.), many middle school teachers that teach untested subject areas, and all high school teachers were not assigned an mSGP.

When the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment replaced the NJASK in the 2014-2015 school year, districts were required to incorporate mSGP scores from the PARCC data that were released. While the requirements for teachers receiving an mSGP remained the same, it is possible that the comprehensive nature of the PARCC assessment (administered in Grades 3-11) will result in a greater percentage of teachers receiving mSGP scores. In light of the recent and ongoing changes made to mSGPs in New Jersey, it is likely that Student Growth Percentiles will continue to be a topic of conversation among local, state, and national education agencies.
The Marzano Causal Teacher Evaluation Model

Robert J. Marzano, the co-founder and current CEO of Marzano Research Laboratory in Denver, Colorado, and Executive Director of the Learning Sciences Marzano Center in Palm Beach Gardens, Florida, began his career in education in 1967 in New York City, where he worked as an English teacher (Marzano, 2013c; Quinn, 2014; Scherer, 2001). The focus of his educational research was and continues to be the qualities of effective instruction, with the primary goal of enhancing student achievement. In 2003, Marzano’s book *What Works in Schools: Translating Research into Action*, was published, in which he identified the following five “school level factors” that impact student achievement: (1) guaranteed and viable curriculum, (2) challenging goals and effective feedback, (3) parent and community involvement, (4) safe and orderly environment, (5) collegiality and professionalism.

These school level factors laid the foundation for several of Marzano’s future publications that he authored and co-authored, identifying best practices in instruction including the following: *Classroom Management that Works* (Marzano, Marzano, & Pickering, 2003), *Classroom Assessment & Grading that Work* (2006), *The Art and Science of Teaching* (2007), and *The Highly Engaged Classroom* (Marzano, Pickering, & Heflebower, 2010). Marzano employed the research-based effective instructional strategies discussed in *What Works in Schools* as a foundation for identifying the characteristics of effective educational leaders and components of effective supervision. Marzano discussed these topics in *School Leadership that Works* (Marzano, McNulty, & Waters, 2005) and *Effective Supervision: Supporting the Art and Science of Teaching* (Marzano, Frontier, & Livingston, 2011). The Marzano Causal Teacher Evaluation Model, a model developed through synthesis of years of research conducted by
Marzano and his colleagues, is used in whole or in part in over 30 states in the United States (Quinn, 2014).

In his article “The Two Purposes of Teacher Evaluation,” Marzano (2012) stated: “An evaluation system that fosters teacher learning will differ from one whose aim is to measure teacher competence.” Throughout his research and publications, Marzano has been consistent in his approach to student achievement, teacher quality, and effective supervision: Marzano continues to assert that effective schools employ specific, personalized objectives for students and faculty and provide targeted feedback to support academic growth and foster teacher development.

Teacher development is achieved through the use of a teacher evaluation system that has three primary characteristics: (1) the system is comprehensive and specific, (2) the system includes a development scale, and (3) the system acknowledges and rewards growth (Marzano, 2012). Marzano’s sentiments are echoed in current publications addressing the need for teacher evaluation systems that support and reward teacher growth rather than those that rank and rate teachers for the purpose of making personnel decisions (Simon, 2012; Collinson et al., 2009; Arneson, 2015).

The Marzano Causal Teacher Evaluation Model (hereafter referred to as the Marzano Model) has been used as an evaluation tool in a variety of states either in part or in whole (Marzano, 2013). Marzano, Frontier, and Livingston (2011) suggested, “The primary purpose of supervision should be the enhancement of teachers’ pedagogical skills, with the ultimate goal of enhancing student achievement” (p. 2). Through the development of the Marzano Model, Robert Marzano provided a framework from which educational leaders could assess teacher quality and provide targeted feedback in an effort to enhance pedagogical skills.
In his book, *The Art and Science of Teaching*, Marzano (2005) asserted that effective teaching involves a combination of science (knowledge of effective classroom strategies and behaviors) and art (knowledge of when to apply the classroom strategies in individual contexts). While there is no mathematical or scientific formula that exists to guarantee teacher effectiveness, Marzano (2003b) developed a framework for identifying characteristics of effective schools and effective teachers. According to Marzano (2003), effective teachers (1) use effective instructional strategies, (2) use effective classroom management strategies, and (3) follow an effective classroom curriculum design. The Marzano Model integrates each of these characteristics into a comprehensive and specific evaluation system. Marzano defined a comprehensive evaluation model as one that “includes all those elements that research has identified as associated with student achievement” and a specific evaluation model as one that “identifies classroom strategies and behaviors at a granular level” (“Marzano, 2012, p. 16).

The Marzano Model is segmented into four different Domains: Classroom Strategies and Behaviors (Domain 1), Planning and Preparing (Domain 2), Reflecting on Teaching (Domain 3), and Collegiality and Professionalism (Domain 4). The four domains include 60 elements: 41 elements in Domain 1, eight elements in Domain 2, five elements in Domain 3, and six elements in Domain 4 (Marzano, 2011).

Marzano, Frontier, and Livingston (2011) argued, “What occurs in the classroom has the most direct causal link to student achievement” (p. 5). Therefore, Domain 1 of the Marzano Model contains the largest number of elements. In an interview published by *Educational Leadership*, Marzano cited eight effective classroom practices that increase student achievement: identifying similarities and differences; summarizing and note taking; receiving reinforcement for effort and recognition for achievement; doing homework and practicing; using nonlinguistic
representations; learning cooperatively; setting objectives and receiving feedback; generating and testing hypotheses; and using cues, questions, and advance organizers (Scherer, 2001). Effective teachers, therefore, utilize each of these classroom practices.

The 41 elements of Domain 1 indicate specific ways in which teachers implement effective classroom practices and are divided into nine Design Questions (DQs): Communicating Learning Goals and Feedback (DQ1), Helping Students Interact with New Knowledge (DQ2), Helping Students Practice and Deepen New Knowledge (DQ3), Helping Students Generate and Test Hypotheses (DQ4), Engaging Students (DQ5), Establishing Rules and Procedures (DQ6), Recognizing Adherence to Rules and Procedures (DQ7), Establishing and Maintaining Effective Relationships with Students (DQ8), and Communicating High Expectations for All Students (DQ9). The Design Questions are framed as questions for teachers and act as categories for assessment within the evaluation model. For example, Marzano (2005) established the following as a question for teachers in *The Art and Science of Teaching*: What will I do to establish and communicate learning goals, track student progress, and celebrate student success?

The nine Design Questions are further broken down into three lesson segments: Lesson Segment Involving Routine Events (DQ1, DQ6); Lesson Segment Addressing Content (DQ2, 3, 4); and Lesson Segment Enacted on the Spot (DQ 5, 7, 8, 9) (Marzano, 2011). A lesson segment, as defined by Marzano (2011) is “an event in the classroom that has a specific purpose and a specific set of teacher behaviors and strategies that are designed to meet that purpose” (p. 34). The lesson segments further compartmentalize the instructional strategies, providing teachers categories on which to focus when planning and preparing lessons.

Domain 2 addresses planning and preparing the following: (1) lessons and units, (2) materials and resources, and (3) special needs of students (Marzano, 2011). The act of planning
and preparing is directly related to classroom behaviors and strategies. If a teacher does not adequately prepare the three categories of activities, she cannot successfully implement the intended strategies (Marzano, Frontier & Livingston, 2011). The tenth and final Design Question from *The Art and Science of Teaching* addresses planning and preparation: What will I do to develop effective lessons organized into a cohesive unit?

According to Marzano, Frontier, and Livingston (2011), Domain 3 (Reflecting on Teaching) “might be thought of as a metacognitive aspect of teacher growth and development” (p. 48). Through use of the Marzano Model, teachers are expected to first evaluate their own professional performance and then develop and implement a professional growth plan that is reflective of areas that are in need of improvement (Marzano, Frontier, & Livingston, 2011).

Finally, Domain 4 (Collegiality and Professionalism) addresses the ways in which teachers promote positive school environment, an ongoing exchange of ideas and strategies in the interest of professional development, and district and school improvement (Marzano, Frontier, & Livingston, 2011). In the book *What Works in Schools*, Marzano (2003) identified collegiality and professionalism as one of the critical attributes of effective schools.

As one of the necessary characteristics of an effective teacher evaluation system is a component that acknowledges and rewards growth, the Marzano Model employs the use of scales to assess teacher effectiveness. The scale is broken into the following developmental levels: Not Using, Beginning, Developing, Applying, and Innovating. At the Not Using level, a strategy is called for but is not used by the teacher; at the Beginning level, the teacher uses the instructional strategy but “with errors and omissions”; at the Developing level, the teacher utilizes the strategy correctly but does not actively monitor whether or not the strategy elicits the desired response from students; at the Applying level, the teacher correctly utilizes the strategy
and monitors the class to ensure that it is having the “desired effect”; finally, at the Innovating level, the teacher goes beyond the criteria for Applying because he adapts the strategy to reflect the needs of all students during instruction (Marzano, 2012, p. 18).

**Common Problems Associated with Teacher Evaluation**

Traditionally, teacher evaluations have served little purpose in aiding professional growth and development. Teachers were observed for approximately one hour of instruction, an extraordinarily disproportionate amount of time when compared to the number of hours a teacher spent instructing students in a given year. The evaluator provided superficial feedback that was specific to the behaviors observed during the lesson, hardly enough data to determine the effectiveness of the teacher. Most often, teachers were rated as Effective, thus maintaining the bureaucratic status quo of teacher evaluation in an attempt to maintain compliance with state evaluation policies (Danielson, 2012; Papay, 2012; Sanders & Horn, 1998). This evaluation process often elicited fear and negativity among the teachers as well (Block, 1992; Conley & Glasman, 2008; Garrett, 2011).

Fear permeated the education profession when high-stakes teacher evaluation was introduced across the United States. These apprehensions were not, however, entirely unfounded. After implementing Impact, a high-stakes teacher evaluation system, 165 teachers were fired from Washington, DC public schools (Dillon, 2011). Proponents of the evaluation system reported that the evaluation system helped teachers to become more effective through collaboration between evaluator and teacher and more detailed and specific feedback from evaluators. Those who opposed the new evaluation system criticized the system for failing to differentiate between poverty-stricken and affluent districts within the District of Columbia. Teachers also claimed that they were being “nitpicked” by the evaluators, two of which were not
administrators in the district, but contracted “lead evaluators.” Teachers feared taking instructional risks for fear of poor evaluation scores and potential dismissal (Dillon, 2011).

The Impact teacher evaluation system is just one example of high-stakes evaluation presently being implemented. Current teacher evaluation practices are reminiscent of the high-stakes testing many policymakers and public officials have criticized. The issue with high-stakes evaluations remains the same: there are legitimate concerns regarding the validity and reliability of the multiple measures from which teachers’ summative ratings and, consequently, employment statuses are derived. Validity refers to whether the scores of the assessment accurately measure the intended trait or characteristic, in this case the quality of the teacher; reliability refers to the consistency of the results yielded from the assessment (Herlihy et al., 2014).

In the case of validity, the evaluation instrument itself (i.e., The Marzano Causal Evaluation Model) is a tool through which evaluators assess teacher effectiveness. While the AchieveNJ Act requires administrators (teacher evaluators) to receive an initial training and subsequent “refresher” trainings on the district-selected evaluation instrument as well as a minimum of two dual observations in order to promote inter-rater reliability, the fact remains that teacher evaluation is subjective (Herlihy et al., 2014; Peterson, 1987, 2000) and, therefore, somewhat lacking in reliability.

Research indicates that outside observers (individuals who have no personal knowledge of the teacher being observed) are more accurate evaluators than principals and evaluators that have regular contact with the teachers whom they are observing, as the latter are more likely to inflate the evaluation ratings (Whitehurst & Chingos, 2014; The New Teacher Project, 2013). Reliability of internal observations is, therefore, a legitimate concern. While accuracy of the
evaluation increases in the cases of highly effective and highly ineffective teachers, evaluators tend to have a difficult time accurately assessing teachers who score in the average ranges (Harris & Sass, 2014).

As noted, a prominent criticism of traditional teacher evaluation was the fact that the process resulted in the large majority of teachers being rated as “satisfactory” or “effective.” Statewide teacher evaluation scores in New Jersey in the 2013-2014 school year, however, did not indicate a significant shift from the trends of the past: 73.9% of New Jersey teachers were rated as Effective; 23.4% of New Jersey teachers were rated as Highly Effective; only 2.5% of New Jersey teachers were rated as Partially Effective; and a mere 0.2% of New Jersey teachers were rated as Ineffective under the 2013-2014 evaluation systems (Mooney, 2015).

From a policy perspective, teacher evaluation policies and procedures are developed by third-party officials: legislators, state department officials, outside vendors and consultants (Hazi, 2014). This era in teacher evaluation has been characterized by federal and state policies and initiatives. Local control is diminishing while more states stake claim over how teachers are evaluated.

The RTTT initiative introduced in 2009 encourages states to implement six core categories of education policies, one of which was teacher evaluation. The initiative required states to adopt evaluation models that incorporated multiple measures to assess teacher effectiveness. As of April 2012, 30 states had adopted evaluation models that included multiple measures, 31 states had adopted evaluation models that included multiple rating categories to assess teacher effectiveness, and 25 states reported that the adopted evaluation policy required evaluators to conduct annual, or summative, evaluations (National Center for Educational Evaluation, 2012). However, whether or not these drastic changes to teacher evaluation have
improved instruction remains to be seen. The emphasis has and continues to be on the implementation of the evaluation system itself and less so on instructional improvement (Weiss, 2014).

In terms of implementing organizational change, the change must be reflected as part of the organizational life as opposed to a response to the most recent evaluation policy (Fullan, 1997). Unfortunately, recent changes in teacher evaluation in New Jersey have more often been the result of the latter than the former. A district must be able to differentiate between and acknowledge the inherent disunion between the technical components of the evaluation system and the factors within the organization that influence the implementation of the evaluation system (Darling-Hammond, 1990; Davis, Pool & Mits-Cash, 2000).

Additionally, one must acknowledge the dual purposes of teacher evaluation (professional growth and quality assurance) and seek to bridge the gap between the two. Policymakers, who are often implementing the required components of teacher evaluation, emphasize the importance of quality assurance from the political perspective. They often argue that teachers are employees of the state, whose salaries are funded by taxpayer dollars. Taxpayers and students alike deserve quality teachers in the classroom, a strong argument for the quality assurance perspective. Conversely, administrators and teachers alike argue that the primary purpose for teacher evaluation is professional development and growth (Danielson, 2001).

Effective Teacher Evaluation
Charlotte Danielson (2012) argues that all teachers, even the most skilled and innovating professionals, can be improved. Teacher improvement occurs from self-reflection and, ideally, an effective evaluation practice. But what does effective teacher evaluation look like? In 1988, the Joint Commission on Standards for Educational Evaluation developed 27 standards that each fall into one of the following four attributes of sound evaluation practices:

- **Propriety** – The evaluation system must protect the rights of all individuals involved in the evaluation process.
- **Utility** – Evaluations must take place in a timely manner and are informative to all parties involved.
- **Feasibility** – Consideration of availability of resources must take place in order for an evaluation practice to be sound.
- **Accuracy** – The evaluation system must accurately reflect the abilities of personnel in order to ensure that sounds judgments and decisions are made as a result of the evaluation process. (Stronge, 2006).

If the Personnel Evaluation Standards are used as recommended, schools will be capable of developing and adopting teacher evaluation systems that “move beyond bureaucratic paperwork to become a critical piece of school reform” (Stronge, 2006, p. 67). A study conducted by Regional Education Laboratory Northeast & Islands and the New Hampshire Department of Education revealed five policy considerations for the implementation of a teacher evaluation system that includes multiple measures. The key ideas presented in the considerations were the need for adequate training and resources for evaluators, adequate time to introduce the instrument in increments, access to annual training on the evaluation instrument for both evaluators and those evaluated, and a positive professional climate to aid in fostering acceptance
of, and ultimately, professional growth as a result of the implementation of the model (Riordan, et al., 2015).

The evaluation process should also include both formative and summative components (Danielson & McGreal, 2000; Hiller, 1986; Howard & McCloskey, 2001). Formative evaluation refers to the components of the evaluation system that support growth and enhance the professional skills of the teacher. Formative assessment may include walk-throughs, peer observations, self-reflections, and goal setting. Summative evaluation is largely based on classroom observations and is intended for the purpose of making personnel decisions about retaining and dismissing teachers, maintaining or rearranging teacher assignments (grade level, content) and, when applicable, differentiated pay values, based on the summative evaluation process (Danielson & McGreal, 2000; Howard & McCloskey, 2001). Evaluation systems that reflect an emphasis on formative evaluation produce higher levels of satisfaction and more reflective practice (Danielson & McGreal, 2000).

Just as teachers are encouraged to differentiate lessons based on student needs, research suggests that a differentiated evaluation system is most effective in improving teacher practice (Danielson, 2001; Glatthorn & Holler, 1987; Howard & McCloskey, 2001). Novice teachers do not possess the same skills as experienced teachers, but professional growth is essential for both subgroups. However, differentiated evaluation processes do not necessarily require different evaluation systems. Effective evaluation is a collaborative process that incorporates ideas and suggestions from both the administrator and the teacher (Danielson, 2001; Danielson & McGreal, 2000; Dennington, 2011; Glatthorn & Holler, 1987; Howard & McCloskey, 2001; Marzano, 2013).
Effective evaluation also incorporates clear and specific expectations of the teacher. These expectations are based on a clearly defined set of standards (Danielson, 2012; Danielson & McGreal, 2000; Shinkfield & Stufflebeam, 1995; Toch, 2008). The Marzano Causal Teacher Evaluation Model, for example, is aligned with the InTASC standards of New Jersey. In order for this alignment with professional standards to have a positive impact on teacher evaluation, teachers must be informed of and trained on the standards and the evaluation model and explicitly instructed on how administrators will assess teacher quality using the model (Danielson & McGreal, 2000; Howard & McCloskey, 2001; Stronge, 2006; Toch, 2008).

In recent years, public policy has reflected a shift from single classroom observations to evaluate teacher effectiveness to more comprehensive teacher evaluation systems that include multiple measures. Measures for consideration have included the following: classroom observations, student growth models, standardized test scores, student surveys, peer observations, and video assessments. The New Jersey Department of Education requires the following three measures are factored into a district’s teacher evaluation system: classroom observations (a minimum of three for tenured and non-tenured teachers), student growth objective (SGO) scores for all teachers, and mSGP scores for qualifying teachers (calculated from standardized test scores). The three measures are weighted and reflect an emphasis on classroom observations or teacher practice (70%-80%) (AchieveNJ, 2014). This fact further signifies the importance for delving into how teachers experience teacher evaluation and the impact this experience has on instruction.

**Teacher Perception of Evaluation**

Much attention has been paid to how students learn and process information, but the research is lacking in understanding how teachers learn and how this learning impacts
instructional practices. Before delving specifically into teacher perception of evaluation, it is necessary to consider how teachers learn and process new information and implement policies regarding educational reform. According to Putman and Borko (2000), cognition, or the act of acquiring new thought, “is (a) situated in particular physical and social contexts; (b) social in nature; and (c) distributed across the individual, other persons, and tools” (p. 4). Putman and Borko are specifically referencing the theory of situated cognition, the idea that the situation in which the individual learns the new information significantly impacts the way the individual processes the information. The ways in which individuals interact with their physical and social environments influence their perception of the new knowledge. This theory is particularly applicable to teachers because teachers are continually experiencing and processing new information in social settings, be they professional learning communities (PLCs), faculty or department meetings, or outside professional development. Teaching is not a profession that promotes solitary learning; in fact, today’s educational culture and climate forbids it. Specific to this study, the ways in which the teacher experiences the learning associated with teacher evaluation will directly impact teacher perception of evaluation and, ultimately, teacher performance.

Teacher evaluation has been characterized as a hierarchical, one-way process through which the administrator offers suggestions to improve teacher practice based on a limited number of classroom observations (Danielson & McGreal, 2000). The administrator is often viewed as a building manager as opposed to an instructional leader, an assumption that leads to mistrust and a lack of administrator credibility in the eyes of the teacher (Danielson & McGreal, 2000; Young & Heichberger, 1975). Low levels of trust between the administrator and the teacher result in a passive evaluation process that is minimally impactful for teachers (Danielson
& McGreal, 2000). Additionally, teachers do not always perceive that the summative evaluations accurately reflect their abilities as educators. In a 2012-2013 study of Arizona school districts implementing an evaluation system consisting of multiple measures similar to that of New Jersey, the Arizona State Department of Education reported that only 39% of teachers surveyed felt that the summative evaluation accurately reflected their abilities; 32 percent indicated that the evaluation did not accurately reflect their abilities, and 30% were undecided (Ruffini et al., 2014).

Teacher perception of evaluation is dependent upon the perceived credibility of the evaluator as well, in that employees are more likely to accept the feedback provided by an evaluator as accurate and make the suggested changes if the employee believes the evaluator has credibility (Kinicki, et al., 2004). Likewise, if a teacher perceives the feedback to be useful, it is more likely that the teacher will utilize evaluator feedback to inform professional judgment and solicit opportunities for professional growth (Tuytens & Devos, 2014). Zimmerman and Deckert-Pelton (2003) identified four key domains that are characteristic of an effective evaluation process: (1) positive interactions between evaluator and teacher, (2) consistent evaluations, (3) principal commitment to professional evaluation, and (4) principal knowledge of pedagogy, content, and evaluation. Teachers look to principals as building leaders, specifically in the area of instructional evaluation. If a teacher believes that a principal is not adept in teacher evaluation, the teacher is far less likely to trust the principal and the integrity of the evaluation process (Zimmerman & Deckert-Pelton, 2003).

In recent years, there has been a shift in the role of the building principal from building manager to instructional leader. The research suggests that strong instructional leadership fosters increased teacher efficacy and, in turn, increased student achievement (Goddard et al., 2015). On
the contrary, if an evaluator’s assessment of a teacher is not aligned with the teacher’s beliefs regarding classroom performance, the teacher’s self-efficacy decreases (Ham, Duyar, & Gumus, 2015). Teacher evaluation can be a powerful tool in improving instruction; however, the teacher must perceive the evaluation system as a useful tool in promoting and increasing student achievement. If this component is not present, it is unlikely that the evaluation system will result in the intended effect. Thus, the body of research on teacher perception of evaluation must be broadened in order to adequately address and combat the aforementioned concerns.
CHAPTER III

METHODOLOGY

Introduction

The purpose of this study is to explore how teachers perceive the influence of the Marzano Causal Teacher Evaluation Model on their professional practices. A phenomenological qualitative research design was most suitable because I sought to describe teachers’ experiences and perceptions of the Marzano Causal Teacher Evaluation Model. Phenomenological qualitative research focuses on the experiences and lived events of humans. Further, the phenomenological study seeks to synthesize commonalities of collective perceptions and experiences related to a particular phenomenon (Saldana, 2011). In this case, the phenomenon being explored was teacher evaluation practices. Semi-structured interviews conducted with selected teachers and school documents pertaining to the evaluation practices and policies in the district were used to examine teachers’ perceptions.

In this chapter, I explain the reasons why I chose to study teachers’ perceptions of the Marzano Causal Teacher Evaluation Model. I then provide a detailed description of the design of the study and the methods used to answer each of the research questions. Finally, I explain how I collected and analyzed the data and how the research procedures employed were validated.

Background

When I began teaching in 2006, there were very few federal or state mandates that addressed teacher evaluation practices. Typically, the school district developed and approved board policies that indicated how often tenured and non-tenured teachers were to be observed. Teacher contracts actually had more influence on evaluation practices than did state or local mandates. As a non-tenured teacher, I was observed three times during my first year of teaching.
The “scale” that was employed was dichotomous, consisting of a checklist of items from which the evaluator determined whether I was “satisfactory” or “unsatisfactory.” The success of this system hinged upon how perceptive the evaluator was and how reflective the teacher was, two variables that could change on any given day. Little attention was paid to the validity or reliability of such a system. However, it was about this time that widespread discussion of teacher evaluation reform began to take center stage in the political arena. Public outcry for “highly qualified teachers” sparked teacher evaluation reform in the state of New Jersey.

By the 2013-2014 school year, per AchieveNJ, school districts were required to implement a state-approved evaluation system. During this time, I was a teacher in a southern New Jersey school district that chose to implement the Danielson model. As a teacher, I was an active participant in this change, attending a variety of training sessions on the model, which included viewing and critiquing videos of teachers in the classroom. These videos were labeled according to the Danielson rubric. Teachers were scored using a scale of 1-4, a practice that most districts emulated in the implementation of the evaluation system. Many, including myself, were leery of this change, a change that seemed to quantify the quality of a teacher, an oxymoron in my book. Yet, teachers’ concerns did nothing to stop this paradigm shift in the field of education.

I watched this change negatively impact teacher morale and incite several veteran teachers to pursue early retirement. I felt that the public was vilifying teachers, most of whom worked tirelessly to positively impact their students’ academic and social growth. I, myself, felt that the new evaluation system implemented in my district did little to improve my performance. When I was told that I should not expect to be a “4” on the Danielson rubric, I immediately felt inadequate. I wholeheartedly agreed that accountability measures should be in place. I was not
at all offended that the process by which teachers could be brought up on tenure charges was streamlined. In my opinion, there were teachers employed in my district that were not effective. In many ways, the new approach to tenure charges sparked a positive change in teacher quality due in large part to the accountability measures that were put in place. I experienced many conflicting experiences and emotions during this seminal year of teacher evaluation reform. During this time, the seed was planted: I wanted to understand why teachers reacted the way they did to this particular change in a field that is accustomed to rapid and widespread change.

During the year I conducted my research, I began my role as an assistant principal of a large middle school in southern New Jersey. This district utilized the evaluation model that I chose to study, The Marzano Causal Teacher Evaluation Model. As I began to adjust to my new role as the evaluator rather than the evaluated, I continued to wonder how teachers perceived the state-approved evaluation systems. However, my new role forced me to ponder another component, the component that was the driving force behind this research study: Do the evaluation systems influence teachers’ instructional decisions? Numerous studies have cited the fact that teacher quality is one of the most significant classroom factors impacting student achievement. I have seen both sides of this debate in my experiences as an evaluator. I have had enriching post conferences, during which teachers reflected upon their lessons and immediately employed my suggestions to improve their practice. I have also seen the other side of the spectrum. Teachers have been deflated after receiving a Developing rating, or outright argumentative after receiving a score with which they did not agree.

Both my experiences as a teacher and an administrator have influenced my opinions of teacher evaluation. As a teacher, I understood the intended purpose but felt that the new evaluation systems missed the mark in several capacities. The changeover was not done
gradually, and the new evaluation systems were drastically different from the approach that had been used for decades. I felt that the summative evaluation rating was not at all reflective of my professional capabilities. It was simply a number that was used to determine my relative value in my district. Additionally, district administrators were not consistent with their evaluations of teachers. Through communicating with other teachers, it was clear that some observers were “harsher” than others when rating specific elements. I was insulted when I was told that I would have to “prove” my professionalism through artifacts that verified my attendance at district events and my participation in professional development.

As an evaluator, my colleagues and I have had many conversations about the ways in which we use the Marzano Model. We have regularly “calibrated” our practice through dual observations and evaluation “refresher” trainings. However, we are still human instruments observing specific elements in each classroom. Our perceptions will inevitably influence the ways in which we evaluate the teachers. I want to help teachers grow professionally, but I do feel that the evaluation process sometimes hinders instead of helps foster professional growth. All of my opinions, while I have identified and recorded them through memo writing, unavoidably influenced my interest in and my approach to the current study.

Of course, the intent of these evaluation systems is to improve teacher instructional practices, but further research must be done in order to uncover whether and how evaluation practices influence teacher quality. These experiences helped to shape the following three research questions that guided this study:

1. How, if at all, has the Marzano Causal Teacher Evaluation Model influenced and informed teachers’ instructional practices?
2. What are teachers’ perceptions of how accurately the Marzano Causal Teacher Evaluation Model reflects and captures their professional performance and capabilities?

3. What are teachers’ perceptions of the influence of the Marzano Causal Teacher Evaluation Model on the relationships with their administrators?

**Design**

I employed the qualitative method for this research design. The qualitative method is focused on the “complexity of interactions expressed in daily life and by the meanings the participants themselves attribute to these interactions” (Marshall & Rossman, 1989, p. 2). Qualitative research is grounded in the concept that the lived experiences of people reveal truths about human existence; these individual “truths” are derived from studying perceptions of the subjects. As this study was designed to explore teachers’ perceptions of the Marzano Model, it was appropriate and relevant to do so through the qualitative research method.

According to Marton (1986), phenomenography is “a research method for mapping the qualitatively different ways in which people experience, conceptualize, perceive, and understand various aspects of, and phenomena in, the world around them” (p. 31). As indicated by the research questions, I was generally seeking to explore how teachers perceived teacher evaluation, the studied phenomenon. In order to do so, I conducted interviews with teachers to uncover how they experienced teacher evaluation and the ways in which these experiences shaped their understandings and perceptions of the tool.

Qualitative research focuses on context, is emergent and evolving rather than prefigured, and fundamentally interpretive (Marshall & Rossman, 1989). In order to better understand teacher evaluation in context, teacher beliefs and opinions about their experiences with the
Marzano Model were collected as data. Their responses were identified as their reality and were treated as such. The qualitative design lent itself to an inductive process of data analysis, a process from which themes and concepts emerged and were subsequently coded. This present study sought to uncover factors that teachers perceived to influence their beliefs and experiences with teacher evaluation. I then sought to uncover how, if at all, these beliefs and experiences contributed to teachers’ overall satisfaction or dissatisfaction with the Marzano Model.

Profile of the Site

For the purpose of this study, a pseudonym, Silver Creek, was used in place of the actual study site. Silver Creek was selected for this study because this school district selected the Marzano Causal Teacher Evaluation Model beginning in the 2013-2014 school year, the year in which the state mandated schools to select a state-approved teacher evaluation model. The district is one of 44 school districts in the state of New Jersey that selected the Marzano Model. According to the 2014-15 New Jersey School Performance Report narrative, Silver Creek High School’s mission is “to develop and maintain a comprehensive educational program that fosters the academic, social, and emotional growth of all students.” The New Jersey State School Performance Report also reveals that the four-year adjusted graduation rate at Silver Creek for the 2014-15 school year was 95%, with a dropout rate of less than 1%. In terms of school-wide achievement, roughly half of the students met or exceeded expectations on the 2014-2015 administration of the PARCC ELA section, and about one-third of the students met or exceeded expectations for the PARCC Mathematics section.

In the 2014-2015 school year, nearly 1,700 students were enrolled in the school building, which houses Grades 9-12. About 15% of those students were identified as students with disabilities (SE), roughly 10% were identified as economically disadvantaged (ED) students, and
less than 1% of students were identified as English Language Learners (ELLs). Forty-nine percent were female, and 51% were male. The ethnic/racial subgroup breakdown reflects that the school has a largely Caucasian population; roughly one quarter of the population identify as being a member of a minority race. The length of the school day is 6 hours and 57 minutes, 5 hours and 36 minutes of which are identified as full time instructional minutes, and 2 hours and 48 minutes are identified as shared time instructional minutes. The student to teacher ratio is 13:1 (NJDOE School Performance Report, 2015).

According to Silver Creek’s Board Policy with regard to evaluation of teachers, the board regulation for teacher evaluation is consistent with the Teacher Effectiveness and Accountability for the Children of New Jersey Act (TEACHNJ and AchieveNJ administrative codes). According to the policy, the board shall annually adopt evaluation rubrics for teachers. The evaluation rubrics shall include four defined annual ratings: Ineffective, Partially Effective, Effective, and Highly Effective. If teachers receive a summative (overall) rating of Ineffective or Partially Effective by the end of a given school year, the teacher will be placed on a corrective action plan (CAP). The policy also indicates that the components of teacher evaluation shall include measures of student achievement and observation data in accordance with the provisions of N.J.A.C. 6A:10-4.2 and N.J.A.C. 6A:10-4.4. Finally, the policy indicates that the teacher practice instrument (the Marzano Model) shall meet the criteria of a teacher practice instrument in accordance with N.J.A.C. 6A:10-6.2.

During the 2015-16 school year, all teachers, both tenured and non-tenured were observed three times. Tenured teachers are required to have three short observations during which the evaluator observes for no less than 20 minutes. Two of the three observations were unannounced (not scheduled in collaboration with the teacher). Non-tenured teachers were
required to have at least two long observations and one short observation. One long observation was announced, while the other two observations were unannounced. When conducting an unannounced observation, the evaluator does not schedule a preconference with the teacher. Only announced observations incorporate a preconference as part of the observation process.

There was one caveat to this policy during the year that I conducted this study. For the 2015-2016 school year, the Silver Creek School District implemented a pilot program, titled Project COACH. Participation in the pilot program was strictly voluntary. Fifteen percent of the teachers in the building signed up to be part of the pilot program that adopted a coaching approach to teacher evaluation. Teachers that chose to participate in the program were assigned one administrator that acted as their coach throughout the school year. Coaches performed shorter (approximately 10 minute), more frequent (a minimum of 6) observations of teachers throughout the school year. Teachers were assigned scores at two points throughout the school year: a midpoint and a final. These scores were utilized to calculate a summative evaluation rating. All other processes were identical to the district’s official evaluation policy.

Teachers are assigned a summative evaluator at the conclusion of each school year. The summative evaluator is responsible for conducting a summative conference with the teacher, during which the teacher and evaluator discuss the teacher’s performance throughout the school year. During this conference, the evaluator also inputs data from the teacher’s SGO. No mSGP scores are a part of the teacher’s evaluations due to the fact that none of the teachers receive an mSGP score. A number ranging from 1-4 is assigned, resulting in the teacher’s summative rating.

In order to maintain compliance with state requirements, the school district requires that evaluators conduct dual observations at least twice a year to further strengthen inter-rater
reliability within the district. However, this does not ensure that all teachers will have a dual observation throughout the course of the school year. Observations are conducted by the chief academic officer, instructional supervisors, and building level administration. Administrators conducting summative conferences have observed the teacher at least two of the three required times. District administrators believe teachers and administrators benefit from the consistency this requirement creates.

**Sampling**

Instructional staff members with at least one complete year of teaching experience from the selected high school were invited to participate in the research by way of a research recruitment letter sent via mass email to each of the classroom teachers employed in the building. Both the building principal and the superintendent approved this email. Fourteen of the 107 classroom teachers volunteered to be interviewed for the study. Teachers who agreed to participate in the study were sent a Demographic Profile Questionnaire that was to be completed and returned to me prior to the scheduled interview. The number of years of experience of each of the teachers varied from moderately experienced (4-9 years) to veteran (10 or more years). The 14 teachers also represented a variety of content areas. It was important to explore the perceptions of teachers with a variety of experience levels and content areas in order to determine if this factor influences the perceptions and experiences of the teachers. No non-tenured teachers volunteered to participate in the study. Two male teachers and 12 female teachers agreed to participate in the study.

Curtis & Gesler (2000) identified effective qualitative purposive sampling techniques. The researchers indicate that purposive sampling is “informed a priori by an existing body of social theory on which research questions may be based’’ (p. 1002).
The following criteria were also included, based on the current literature on teacher evaluation and recommended qualitative research practices:

- Representation from different levels of experience (only teachers who had at least one complete year teaching under the Marzano Model were considered)
- Representation from different age groups
- Representation from a variety of grade levels.

It was important to explore the perceptions of teachers that represented each of the aforementioned populations in order to determine if these factors influence the perceptions and experiences of the teachers. As noted in the limitations of the study, the composition of the sample of the study was not completely reflective of the composition of the teachers employed in the district. However, due to the fact that each of the criteria that were developed from the literature review were met, the sample population was appropriate for the current study.

Table 1 details the demographic profile information for each of the participants in the study. The average age for the participants was 39.9 years old.

Table 1

*Summary of Demographic Profile Information for Each Teacher*

<table>
<thead>
<tr>
<th>Teacher #</th>
<th>Sex</th>
<th>Content Area</th>
<th>Number of Years Teaching</th>
<th>In Person or Phone Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Female</td>
<td>Science</td>
<td>27</td>
<td>In Person</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>Science</td>
<td>16</td>
<td>Phone</td>
</tr>
<tr>
<td>3</td>
<td>Female</td>
<td>World Languages</td>
<td>14</td>
<td>Phone</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>Math</td>
<td>6</td>
<td>In Person</td>
</tr>
<tr>
<td>5</td>
<td>Female</td>
<td>Technology</td>
<td>21</td>
<td>Phone</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>6</td>
<td>Female</td>
<td>History</td>
<td>10</td>
<td>Phone</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>Science</td>
<td>11</td>
<td>In Person</td>
</tr>
<tr>
<td>8</td>
<td>Female</td>
<td>English</td>
<td>14</td>
<td>Phone</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>History</td>
<td>21</td>
<td>In Person</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>World Languages</td>
<td>12</td>
<td>Phone</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>Math</td>
<td>21</td>
<td>Phone</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>Science</td>
<td>15</td>
<td>Phone</td>
</tr>
<tr>
<td>13</td>
<td>Female</td>
<td>World Languages</td>
<td>5</td>
<td>Phone</td>
</tr>
<tr>
<td>14</td>
<td>Female</td>
<td>English</td>
<td>15</td>
<td>In Person</td>
</tr>
</tbody>
</table>

**Data Collection**

Each of the 14 selected participants was interviewed. The semi-structured interviews were approximately 35-45 minutes in length and were recorded using a recording device. Interviews were conducted both in person and over the phone, depending upon the preference of the participant. I transcribed all interviews within one week of the actual interview. The interview protocol helped me to learn about the experiences and perceptions of the teachers regarding teacher evaluation practices, with a specific focus on the ways in which teachers believed that these practices influenced instructional decision-making.

I also completed a document analysis of public records, specifically board policies relating to teacher evaluation in Silver Creek as well as the Silver Creek school website. I also referenced the 2014-2015 Silver Creek School Performance Report as reported by the New Jersey State Department of Education. Artifacts of any professional development flyers, meeting agendas for department or faculty meetings, and training materials relating to teacher evaluation
were also examined for the purpose of gaining a better understanding of how the school handles professional development for teacher evaluation, if at all. As previously noted, all of the informants and schools were assigned pseudonyms to protect the anonymity of the participants and the school community.

Table 2 provides an overview of the data collection procedures. Research Question 1 and its correlating sub-questions are identified and described below as an example.

Table 2

*Overview of the Procedures Used to Collect Data*

<table>
<thead>
<tr>
<th>Sub-questions</th>
<th>Data Source</th>
<th>Objective(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do teachers experience and describe the teacher evaluation practices in the Silver Creek High School?</td>
<td>Interviews, Document analysis of the school’s published policies</td>
<td>To understand how teachers experience and describe the teacher evaluation practices in Silver Creek.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To understand what the school publishes on the district website and approves as board policies in collaboration and agreement with community stakeholders and state mandates.</td>
</tr>
<tr>
<td>How do teachers explain the influence of the Marzano Model, if any, on professional growth?</td>
<td>Interviews</td>
<td>To identify and analyze how teachers explain the influence of the Marzano Model on professional growth.</td>
</tr>
<tr>
<td>What components of the Marzano Model, if any, have hindered teachers’ professional growth?</td>
<td>Interviews</td>
<td>To understand which, if any, components of the Marzano Model have hindered teachers’ growth as practitioners.</td>
</tr>
</tbody>
</table>
Approval to conduct this research study was obtained from the superintendent of Silver Creek High School as well as the building principal. Once I received approval from Seton Hall University’s Institutional Review Board (IRB), I began to conduct the research and interviews at the study site. Before conducting the interviews, I assembled an expert panel that consisted of two experienced administrators and two veteran teachers to field-test the interview questions. This panel was organized and assembled to provide feedback on the interview questions prior to conducting the interviews. The feedback solicited was utilized to revise and edit the interview questions to reflect increased clarity and transparency. No member of the expert panel participated in the research.

The data were collected from one-on-one interviews that were approximately 35-45 minutes in length. All interviews were audio-recorded, using a recording device. All interviews were then transcribed within one week of the actual interview. All interviews were conducted either in the Silver Creek High School building or over the phone during a time frame that was mutually agreed upon. In order to avoid external influences, it was important that the interviews were conducted in a private setting where the researcher and participant could interact independently of other people in the building.

In order to provide structure but also allow for opportunities for me to delve more deeply into participant responses, all interviews were semi-structured, leaving opportunity for follow-up questions where necessary. Semi-structured interviews provide the researcher with a focus for the duration of the interview but allow for the researcher to ask clarifying follow-up questions of the participant when necessary. Semi-structured interviews, therefore, may vary significantly from participant to participant while still providing the researcher with data that are directly related to the research questions (Miles & Gilbert, 2004). Semi-structured interviews are ideal
for exploring participant perceptions due to the fact that semi-structured interviews simultaneously provide a general framework and versatility. The researcher can uncover equally valuable conclusions when analyzing the contradictions in participant responses in addition to the similarities uncovered (Miles & Gilbert, 2004). Table 3 identifies sample interview questions and the correlating theoretical framework that relates to each.

Table 3

*Sample Interview Questions Raised by Theoretical Framework*

<table>
<thead>
<tr>
<th>Interview Question</th>
<th>Theoretical Framework</th>
<th>Theorist(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In what ways, if any, do you think that your approach to lesson planning has been influenced by the district’s use of the Marzano Model?</td>
<td>Reflective Practice Theory</td>
<td>Dewey, Borton, Kolb</td>
</tr>
<tr>
<td>How, if at all, has the Marzano Model improved the quality of professional conversations between you and your administrators?</td>
<td>Reflective Practice Theory</td>
<td>Dewey, Borton, Kolb</td>
</tr>
<tr>
<td>How, if at all, has the Marzano Model helped to define the expectations of your administrators with respect to classroom performance?</td>
<td>Adult Learning Theory</td>
<td>Habermas, Mezirow</td>
</tr>
<tr>
<td>What are your thoughts about the accuracy of the summative evaluation ratings (Highly Effective, Effective, Partially Effective, Ineffective) that you’ve earned under the Marzano Model?</td>
<td>Social Cognitive Theory, Social Cognitive Career Theory</td>
<td>Bandura</td>
</tr>
</tbody>
</table>

Field notes were also taken during each interview. These notes provided me with an opportunity to record observations that were not directly related to what was being stated during the interview and later transcribed. I recorded the physical appearance, mannerisms, and tone of
voice, body language, particular gestures, and notable eye movements of the participant, when
the interview was conducted in person. When interviews were conducted over the phone, I
recorded the participants’ tones when responding to each of the questions, significant pauses
during the conversation, and any verbal gestures that were audible during the interview. These
descriptions helped me to accurately record and analyze the participant’s responses while
increasing the likelihood that I would correctly interpret the participant’s responses.

At the conclusion of each interview, I engaged in member checking, a process through
which the researcher allows the participant to review the information that was recorded and make
any adjustments or additions to the data. I also allowed each participant to review the transcript
prior to data analysis. This provided participants an additional opportunity to correct any errors I
may have made during transcription. Member checking is an added component that increases
the validity of the study (Creswell & Miller, 2000). One of the 14 participants responded with
corrections to the interview transcript, including more concise responses in lieu of her initial
responses. She did not change her stance on the interview questions but rather revised her
responses to reflect a more direct response to the question. The participant’s reviewed transcript
was used throughout data analysis.

All documents obtained from the research are kept in a locked filing cabinet. These items
include flash drives, transcripts, questionnaires, and all other printed materials. To protect the
anonymity of each of the participants, all audio-recordings were deleted upon completion and
participant approval of the interview transcripts. All data will be kept for a period of three years.
After that, all of the aforementioned research materials will be destroyed.
Data Analysis

I recorded specific descriptors for the purpose of exploring potential themes during data analysis. The following descriptors were recorded and eventually coded for each participant: approximate age (a range was offered in an effort to protect the anonymity of the participants), number of years teaching (again, a range was provided), number of years teaching in the current district, and gender.

I relied on teachers to explain and describe their perceptions of the Marzano Causal Teacher Evaluation Model and its perceived influence, if any, on teacher instructional practices. The research questions provided a framework for focus and understanding. The data were then coded. After completing the initial reading, open coding, and focused coding, I developed a list of “super codes.” Themes were identified after organizing, categorizing, and analyzing the “super codes.”

The multi-step data analysis process is detailed below:

1. I read all documents gathered during data collection.
2. I read and completed initial open coding of the data.
3. I performed focused coding. All coded data were recorded using the software program Dedoose in order to efficiently organize the data.
4. Participants were divided based on the following descriptors: approximate age (a range was provided in an effort to protect the anonymity of the participants), number of years teaching (again, a range was provided), number of years teaching in the current district, and gender. Coded data were then separated into “super codes,” or categories that I identified.
5. I developed themes from the codes that were identified during the coding process. A priori (pre-set) codes addressing teacher perception of teacher evaluation and its influence on instructional practices were used as well as codes derived from an inductive review of the data collected (emergent codes). Table 4 identifies all of the codes developed from the coding process.

6. I composed an in-depth interpretation and analysis of each of the themes identified. The interpretation and analysis is included in Chapter IV.

Table 4

*Preliminary Set of Data-Driven Codes*

<table>
<thead>
<tr>
<th>Accountability</th>
<th>Communication</th>
<th>Feedback</th>
<th>Objectivity</th>
<th>Reflective Practice (Pre-set)</th>
<th>Teacher Perceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applying</td>
<td>Compliance</td>
<td>Formative Assessment</td>
<td>Professional Conversations</td>
<td>School Culture</td>
<td>Tenure</td>
</tr>
<tr>
<td>Beginning</td>
<td>Credibility</td>
<td>Highly Effective</td>
<td>Professional Growth (Pre-set)</td>
<td>Student Achievement</td>
<td>Training</td>
</tr>
<tr>
<td>Change</td>
<td>Design Questions</td>
<td>Innovating</td>
<td>Professional Relationships (Pre-set)</td>
<td>Subjectivity</td>
<td>Trust</td>
</tr>
<tr>
<td>Clarity in Expectations (Pre-set)</td>
<td>Developing</td>
<td>Inter-rater Reliability</td>
<td>Professional Risk-taking</td>
<td>Summative Evaluation</td>
<td></td>
</tr>
<tr>
<td>Coaching</td>
<td>Effective</td>
<td>iObservation</td>
<td>Proficiency Scale</td>
<td>Teacher Efficacy</td>
<td></td>
</tr>
</tbody>
</table>

**Protection of Human Subjects and Ethical Considerations**

It is a most basic premise for all professions and academic endeavors that involve human subjects to “do no harm” to the patient/participant. The University of California, Los Angeles (UCLA) identified the following six key principles of ethical consideration when conducting research on human subjects:
Throughout my research study, I employed each of these principles. The study holds value due to the fact that it has the potential to inspire positive change regarding effective teacher evaluation. The study was scientifically and methodologically sound. Fair subject selection was employed through use of purposive sampling procedures. Most importantly, throughout the study, all participants that opted to participate in the voluntary study were treated with respect through the use of pseudonyms to protect anonymity and member checking to ensure that I did not misinterpret or misrepresent the intentions or opinions of the participants. Due to the fact that these ethically sound practices were employed, there was a favorable risk-benefit ratio, as participants were exposed to minimal, if any, risk for participating in the study.

Validity and Reliability

Roberts, Priest, and Traynor (2006) define reliability as “how far a particular test, procedure, or tool . . . will produce similar results in different circumstances, assuming nothing else has changed” (p. 41). In qualitative research, trustworthiness of the researcher is essential in establishing reliability (Golafshani, 2003). If the researcher does not accurately report the results of the study, the study and the results of the study are inherently flawed. Some argue that qualitative research is better described as having “dependability” as opposed to reliability (Lincoln & Guba, 1985, p. 300). In order to combat this reality, good quality research employs a
sound research design that incorporates a series of effective steps in qualitative research. For example, I transcribed each of the interviews after the fact but utilized an audio-recording device while conducting the interviews. I also recorded notes while interviewing each of the participants that included facial expressions, physical appearances, reactions, and tones of each of the participants. I also listened for tone when reviewing the audio recording during transcription. These processes not only allowed me to review the data multiple times, but also allowed for me to include significant nonverbal aspects of communication, a component that increased the reliability of the research process (Roberts, Priest, & Traynor, 2006).

The use of Dedoose, a web application frequently utilized in qualitative research, was used to increase the reliability of the research process as well. After the initial transcription of each of the interviews, I uploaded all of the transcripts to the program and engaged in the coding process through mining the data and highlighting significant excerpts. A priori codes were identified through a comprehensive literature review prior to conducting the interview; inductive, or emergent, codes were identified during the coding and analysis process. Dedoose was particularly helpful in identifying excerpts that aligned with a priori codes and sifting through the data to identify inductive codes from the data collected.

Lub (2015) defines validity as “the degree to which the indicators or variables of a research concept are made measurable, and accurately represent that concept” (p. 2). For the purpose of this study, I was exploring teacher perception of the Marzano Causal Teacher Evaluation Model. The indicators of teacher perception of this evaluation instrument were derived from semi-structured interviews with the participants. Themes were identified through the categorization of the data, and the responses of the teachers. If a future study were to follow
the same research design, it is likely that comparable codes and themes would be identified due to the fact that the interview protocol and a priori codes were derived from the existing literature.

One recommended practice in qualitative research is the assemblage of an expert panel or focus group in order to test the interview protocol prior to the study. In order to increase the likelihood that the research design and data gathered from the study were valid, I assembled an expert panel to review and critique the interview questions prepared for the semi-structured interviews. I also employed the qualitative method of member checking in order to ensure that the data gathered were accurate and valid in the eyes of the participant. Validity is also associated with the rigor and quality of the research design (Golfashani, 2003). Through the incorporation of recommended practices in qualitative research, the rigor and quality of the study were inherently improved.

**Role of the Researcher and Researcher Bias**

Those who are skeptical of qualitative research designs have often cited the bias of the researcher as a flaw in its implementation. In qualitative studies, the researcher is considered to be the research instrument. In order to increase the reliability of the research design and the validity of the results, it is imperative that the researcher clarifies her role in the research process and follows a plan of inquiry (Sanjari et al., 2014).

In phenomenological qualitative research, researchers are tasked with transforming the data into the “lived experience” (Sanjari et al., 2014, p. 2). Both researchers and participants have a subjective influence on the outcome of the research due to the fact that the researcher’s perception may influence the interpretation of the participant’s responses, thus impacting the reliability of the data retrieved (Darawsheh, 2014). Of course, the perspectives, biases, experiences, and interpretations of the qualitative researcher cannot be overlooked. The opposite
is actually the case; the researcher must clearly identify and define the aforementioned personal components and develop strategies through which these components will minimally interfere with the study. This process, known as reflexivity, allows for the researcher to reflect upon their cultural background, thoughts, actions, emotions, assumptions, and unconscious responses, while identifying how these factors may influence the research process and findings (Darawsheh, 2014).

For example, I had to acknowledge and account for the fact that I harbored my own opinions and judgments about teacher evaluation from my experiences as a teacher. I knew that I would inevitably have my own thoughts about the responses of the participants, some of which may not have been in agreement with their responses. In order to account for these factors, I conducted memos during the research process. This strategy allowed me to not only acknowledge, but also record and analyze my biases throughout the research process.

Mears (2009) identified the ways in which background knowledge can either increase or decrease the likelihood of researcher bias. Prior to conducting the research for the present study, I was a teacher for ten years. I acknowledged the fact that I had ten years of teaching experience that had shaped my personal views of teacher evaluation. I conducted a comprehensive literature review on the subject of teacher perception of evaluation. Throughout this process, I gained insight and perspectives on the history of teacher evaluation, best practices in implementation of teacher evaluation, and the common problems associated with teacher evaluation. I believe that my experience with and knowledge of teacher evaluation practices of both the Danielson and Marzano Models helped to decrease researcher bias due to the fact that I had been exposed to teacher evaluation from the perspective of both the teacher and the administrator.
I was also an assistant principal in a different school building in the same district as the school that served as my research site. While I did not know any of the participants, I worked with their instructional supervisors and knew their principal, all of whom evaluated the teachers studied. I was, technically speaking, a stranger in my role as the researcher, but participants were made aware that I had professional ties to their immediate supervisors. In order to limit the impact that my professional status had on the study, I assured all participants that they would be granted complete autonomy and that all confidentiality measures would be employed in order to protect their identities. Additionally, I ensured that all participants signed an informed consent form. This form further indicated that participants would be aware of the information that was gathered and reported (Sanjari et al., 2014). These additional steps helped to build trust with the participants and increased the likelihood that they would provide open and honest responses to the interview questions posed.

As the qualitative researcher is the instrument from which the data are obtained, processed, and interpreted, it is essential that the interviews conducted by the researcher are thoughtfully and deliberately planned and executed. For the purpose of this study, I employed the method of designing a semi-structured interview. I designed study-specific, open-ended questions that allowed for respondents to add depth to their responses when they felt it was necessary to do so. These questions were piloted by an expert panel, a strategy that increased the reliability of the interview protocol and the overall rigor of the study (Chenail, 2011). I then added follow-up questions when necessary. These follow-up questions were not leading but rather probing for further detail from the respondent.

No research study, either qualitative or quantitative in design, is entirely free of subjectivity. However, the researcher can decrease the amount of subjectivity in a qualitative
study through truthful recognition of potential personal biases. As indicated in the Background section of Chapter III, I recognized and considered my personal biases throughout the research process through frequent deliberate and meaningful reflection. My intention throughout this study was to explore the ways in which teachers perceived the influence of teacher evaluation practices. After reviewing recommended practices in qualitative research as well as methodological exemplars in the field of qualitative research, I was able to design a research study that reflected a primarily objective exploration of the topic.

**Summary**

Chapter III provided an extensive review of the current study, specifically an in-depth discussion of the background of the study, sampling methods, a profile of the research site and research participants, and data collection and data analysis methods. I also discussed the validity and reliability of the study and identified the role of the researcher and the biases that I recognized throughout the research process. All of this information was described in an effort to illustrate how I effectively addressed each of the proposed research questions.
CHAPTER IV

FINDINGS

Introduction

Chapter IV presents the significant findings and an in-depth analysis of these findings. The first section provides an overview of the purpose and context of the study. The second section addresses each of the three research questions that guided the study and identifies the themes that emerged relating to each of the research questions. The final section presents a summary of the findings presented.

Effective instruction is essential in fostering student growth. Teachers are encouraged to be reflective practitioners, to continually seek methods to implement research-based effective instructional strategies. The Marzano Causal Teacher Evaluation model identifies 41 elements that describe effective instructional practices. The purpose of this study was to explore how teachers perceive the influence of the Marzano Causal Teacher Evaluation Model on their professional practices. I sought to determine if teachers perceived the Marzano Causal Teacher Evaluation Model as a tool to foster professional growth.

Silver Creek High School was selected as the research site due to the fact that the teaching staff at Silver Creek had been utilizing the Marzano Causal Teacher Evaluation model for three complete years prior to the study. The school district implemented the Marzano Causal Teacher Evaluation Model in compliance with state requirements reflected in AchieveNJ, initially enacted during the 2013-2014 school year. All of the teachers that participated in the study had experienced both the district’s former evaluation tool and the current Marzano Causal Teacher Evaluation Model. The participants provided perspectives on both of these distinctly different evaluation tools.
Fourteen Silver Creek High School teachers participated in a semi-structured interview composed of questions specifically designed to address the three research questions that directed the focus of the study. These interview questions are the following:

1. How, if at all, has the Marzano Causal Teacher Evaluation Model influenced and informed teachers’ instructional practices?

2. What are teachers’ perceptions of how accurately the Marzano Causal Teacher Evaluation Model reflects and captures their professional performance and capabilities?

3. What are teachers’ perceptions of the influence of the Marzano Causal Teacher Evaluation Model on the relationships with their administrators?

These research questions were developed as a framework from which to determine significant themes regarding teacher perception of the Marzano Causal Teacher Evaluation model. Following is a presentation and analysis of the themes derived from each of the three research questions.

**Summary of Findings for Research Question 1**

Participant responses revealed that teachers are conflicted about the impact the model has had on their professional practices. While teachers believe that the specificity of the model has helped them to incorporate a variety of research-based instructional strategies, the majority of the participants still identify compliance as a primary reason for changing their approach to instruction. Lesson planning, in particular, is perceived as an unnecessary exercise in compliance. The majority of the teachers did not believe that the new lesson planning requirements aided in improving their practice. Teachers were more likely to try to “fit” the model into their teaching style as opposed to adapting their instruction to reflect the components
of the model. Most of the participants believed that they were already practicing most of the instructional strategies listed in the model. These responses beg the question: How much has instruction truly changed as a result of the change in the evaluation instrument and process?

Teachers did credit the Marzano Model with increasing their knowledge of and attention to formative assessment. Teachers believed that they were more adept in incorporating meaningful formative assessments to check for student understanding as a result of the Marzano Model. They believed that this had a positive impact on student achievement.

Additionally, teachers were supportive of the incorporation of learning goals. The majority of the teachers felt that student learning goals helped to improve their curriculums and streamline their instruction. However, the teachers were not as positive about the student self-assessments, or scales, that the Marzano Model recommends. Teachers felt that the scales were more appropriate for elementary students and were not developmentally appropriate for high school students.

In general, teachers accepted the model as a sound instrument from which they could improve some components of their instruction. Many of the participants preferred the Marzano Model to the previous evaluation tool due to the Marzano Model’s specific framework and documented examples of evidence. Teachers also mentioned the DQs several times, indicating that they felt that they were more cognizant of requiring students to delve more deeply into the content and emphasizing depth over breadth. Yet, most of the teachers did not accept the model in totality but rather interpreted the framework as a list of suggested practices that did not necessarily apply to every lesson.
Themes: Research Question 1

Research Question 1: How, if at all, has the Marzano Causal Teacher Evaluation Model influenced and informed teachers’ instructional practices?

Interview Questions 1-4 were designed to address Research Question 1 through uncovering teachers’ perceptions of the practical application of the recommended practices to which the Marzano Causal Teacher Evaluation Model subscribes. Specifically, teachers were asked to reflect upon their thoughts regarding the instructional and pedagogical emphases of the model as well as how the implementation of the model has influenced their approach to lesson planning.

Perceived Instructional and Pedagogical Emphases of the Marzano Causal Teacher Evaluation Model as Implemented in Silver Creek High School

The participant responses suggested that professional reflection, instructional strategies, and individualized instruction were the three major themes for teachers’ perceptions of the instructional and pedagogical emphases of the Marzano Model as implemented by the Silver Creek School District.

Thirteen of the 14 teachers interviewed mentioned professional reflection as a major emphasis of the Marzano Model. In response to her perception of the emphasis of the model, Teacher 3 responded, “It’s a great way for us to reflect and remind us of what we are incorporating into our practice. I think it’s important for us to be mindful of these things and not to get too, I would say, overwhelmed. But, it’s a really good reflection piece.”

Of the 13 teachers that discussed professional reflection, 11 teachers suggested that the emphasis on professional reflection has instilled a desire to grow professionally, one teacher was indecisive as to whether or not the Marzano Model had helped her to grow professionally, and one teacher indicated that he has not noticed any professional growth since the implementation
of Marzano Model. Teacher 2 identified a connection between professional reflection and growth and the negativity associated with the district’s implementation of the Marzano Model:

My general experience has been that the people that are complaining the most about their evaluations or the most indignant about their evaluation tend to be the ones that are the least reflective, and I think it would be helpful, if at all possible, to try and guide those teachers into self-reflection.

While the majority of the participants noted professional reflection as an emphasis of the Marzano Model, it should be noted that only one of the 14 teachers interviewed indicated that the model has not changed his teaching practices. All of the remaining 12 participants attributed professional growth to the district’s implementation of the Marzano Model since the year that it was introduced.

Teacher 9 expressed skepticism towards the model when asked about the impact the Marzano Model has had on his approach to teaching:

I think it’s just giving us little key words that administrators say to us. I don’t know if it’s necessarily the right thing to do, or the perfect way to teach. It’s just what’s being looked for at this particular moment. I’ve been around in education long enough to see many different flavors of the month. I think this may be another prepackaged, let’s try to fit education in this little box here. We’ll see where this goes three, four, five years from now.

When asked about the perceived weaknesses of the model, several of the participants indicated that the model itself is “overwhelming.” Three participants attributed the identified overwhelming nature of the model to the 41 possible elements, or instructional strategies, that are listed in Domain 1 of the Marzano framework. Instructional strategies were mentioned 11 times
total in seven participant interviews. Interestingly, of these 11 comments, six remarks reflected a positive opinion of the inclusion of the 41 instructional strategies and five reflected a negative opinion. Two teachers offered conflicting opinions regarding the instructional strategies. Teacher 12 is one such example. Teacher 12 stated, “Some of these elements are a lot stronger for certain subjects than they are for other subjects. I’ve learned that some work better for my teaching than others.” However, when asked about how she has utilized the model to inform her instructional practice, she indicated, “[The Marzano Model] provides better feedback for your teaching and what you’re doing. It allows you to look at your teaching style and see if maybe you should be doing a little bit more of another style.” Teacher 12 expressed concern that the model was too narrow for certain content areas but did identify that the inclusion of the instructional strategies has aided in identifying ways that she can vary her approach to teaching her content.

A final theme that emerged relating to the instructional and pedagogical emphases is individualized instruction. Seven of the 14 participants indicated that the Marzano Evaluation Model emphasizes individualized, or differentiated, instruction. Marzano categorizes students of varying abilities within the framework, identifying exceptional student populations as “high-expectancy learners” and “low-expectancy learners.” Teacher 8 addressed both types of learners in her response:

Sometimes you forget about the importance of reaching those lower-level kids. And I think this makes you focus on them and that you’re using strategies to reach them, not just your typical teaching techniques. As a special education teacher, I think that’s great to have all teachers focus on. When they don’t get it, let’s try to figure out how we can help them to get it and try to reach these students that are at risk . . . while still addressing
the higher level kids and making sure that they can still extend their knowledge and reach
deep levels.

**Approach to Lesson Planning**

The participant responses suggested that compliance and student outcomes were the two
themes that emerged relating to teachers’ approach to lesson planning since the Silver Creek
school district implemented the Marzano Model.

With regard to compliance, five of the 14 participants indicated that they have changed
their approach to lesson planning for no other reason than to meet the district requirements with
respect to lesson planning. When asked about lesson planning requirements, the Chief Academic
Officer of the Silver Creek School District stated, “Lesson plans are due on Monday at 8:00 a.m.
For Special Education teachers (in an ICS classroom), lesson plans are due on Tuesday at 8:00
a.m. with specific modifications and accommodations noted.” She also referenced a template
(see appendix) that identifies the required components of teacher lesson plans. In a list form,
these are as follows: unit title, appropriate standards, learning goal, day of the unit, learning
objective(s), learning activities: (including elements from DQ 2, 3, and 4 that will be utilized),
design question(s), learning assignments, and linked assessments.

Teacher 14 expressed her concern regarding the district’s expectations of lesson plans:
“I’ll be honest with you, the lesson plans that I submit are more for administration because they
don’t help me. I’d rather have bullet points of what I’m going to do, what quotes I’m going to
pull out, that’s all I need.”

Teacher 3 specifically voiced her frustration with the fact that teachers are required to
include specific elements that they are addressing in their lesson plans: “I think it can be a
frustrating thing. It’s very frustrating when we have to write [the elements] down. We have
been applying them, but we just have to find what elements we are applying in each lesson.” Her sentiments were echoed by the other four teachers, who cited compliance as the primary reason for adapting their lesson plans.

The nine remaining participants indicated that they are more focused on student outcomes as a result of the Marzano Model. Four teachers expressed that the Design Questions (DQs) in Domain 1 of the Marzano Model had a significant impact on their approach to lesson planning, and, subsequently, their focus on student achievement.

Teacher 11 emphasized the importance of DQs in her current approach to lesson planning: “In my lesson plans, I am very serious about [the DQs] . . . because I feel like students need to tell me what they’ve learned because if I’ve taught an entire unit and they get a 0, what did I accomplish? What did they learn?”

As part of the professional development provided in recent years related to lesson planning, Silver Creek has highlighted the importance of the “KUD” approach to lesson planning, which prompts teachers to consider what they want students to be able to Know, Understand, and Do. More than half of the teachers that participated in the study suggested that the Marzano Model has heightened their awareness of the importance of being cognizant of what they want students to know, understand, and do. Teacher 10 described the impact the Marzano Model has had on her approach to lesson planning:

For me personally, what I’ve really gotten from Marzano, what has made me a better teacher over the past few years, is truly understanding why I’m teaching what I’m teaching. So, although we may have a given curriculum, I am now forced to really understand, even though this is the given curriculum, why am I teaching it? What do I need students to be able to walk away and understand? How am I best going to be able to
implement different strategies, different techniques, to allow them to understand what that is?

According to the Marzano Teacher Evaluation Model, student achievement is the “nonnegotiable goal for instruction.” The majority of the participants agreed with Marzano’s prescribed purpose of instruction.

**Summary of Findings for Research Question 2**

Of the three research questions that guided the study, participants reacted the most strongly to the interview questions that addressed Research Question 2. Teachers developed a negative opinion towards the Marzano proficiency scale ratings and the summative evaluation ratings. By and large, participants expressed their desire to be Innovating and Highly Effective, but perceived these ratings to be largely unattainable. Teachers associated Applying and Effective ratings as failures due to the fact that they could not earn the highest rating on the scale. It was not enough for participants to be Applying in an observation or Effective at the end of the year. While they were told that they were doing their jobs well by the administrators, the majority of the participants felt that they were destined for failure because they could never achieve the “A” grade that they were striving for. As a result, teachers reported a drop in morale and a decreased sense of efficacy after the Marzano Model was introduced.

Teachers also felt that administrators were not clear when they explained what Innovating “looked like” in the classroom. A few participants indicated that they must do something “new” to earn an Innovating rating, but the rest of the participants were not able to communicate a district definition for Innovating. Along the lines of clarity in expectations, most teachers identified inter-rater reliability as a weakness of the Marzano Model in Silver Creek High
School. The fact that teachers could not articulate district expectations with respect to the Marzano Model reflects ambiguity and uncertainty from the teacher perspective.

Finally, teachers expressed a concern regarding the number of observations that are performed throughout the school year. Several teachers joked that they did not want administrators in their rooms every day but said that they would appreciate an evaluation tool that reflected a more accurate picture of their professional capabilities. Teachers believed that a total of 80 minutes in their classrooms did not accurately convey their professional skills and efforts throughout the course of the school year.

**Themes: Research Question 2**

Research Question 2: What are teachers’ perceptions of how accurately the Marzano Causal Teacher Evaluation Model reflects and captures their professional performance and capabilities?

Interview Questions 5-7 were designed to address Research Question 2 through uncovering teachers’ perceptions of the accuracy and validity of the ways in which they are evaluated under the Marzano Teacher Evaluation Model. Themes that emerged as a result of asking the interview questions related to Research Question 2 were inter-rater reliability, clarity of expectations, depth and breadth of the evaluation, constructive and useful feedback, and perceptions of proficiency scales and ratings.

**Inter-rater Reliability**

Five of the 14 participants cited inter-rater reliability as a concern associated with the perceived validity and accuracy of teacher evaluation practices in Silver Creek High School. Even though each of these five teachers acknowledged the fact that the district policies and procedures related to teacher evaluation were implemented to protect teachers from administrator
bias, the conversations with the participants revealed that a portion of the teachers are still concerned about the subjectivity of the evaluation model. Teacher 7 stated the following:

So, when you have two different human beings, who, even though they are supposed to be looking for the same things, they are two different people. It doesn’t matter how you dice that up. If they have a preference or a bias towards a certain person, it is really difficult with that human element when you have people coming in and observing you on these elements.

Teacher 10 provided a specific anecdote to shed light on the issue of inter-rater reliability:

I’ve been evaluated two years in a row in the same lesson. In one, I got Applying, and in another, I got Developing, and I did the same thing. That bothers me. Because why did you give me Applying, and why did you give me Developing? And I mean the person who gave me Applying was not the one that gave me that the second time. I’m like, “Why did you downgrade my lesson plan?”

The aforementioned example reveals the flaw in any human-operated assessment tool: you cannot remove subjectivity entirely. One participant applauded the efforts of the administration in their attempts to develop a more objective evaluation process through the implementation of dual observations; but in spite of the district’s efforts, some teachers are still not confident in the inter-rater reliability with respect to the Marzano Causal Evaluation Model.

Clarity of Expectations

Closely linked to inter-rater reliability is the teachers’ perceived level of clarity of the district’s expectations of teachers with respect to teacher evaluation. Eight of the 14 participants mentioned clarity of expectations when answering the interview questions. Teachers were split
down the middle with respect to whether they felt that the district’s expectations were positively or negatively impacted by the Marzano Model. Teacher 9 directly stated, “If the teachers know what they are going to be scored on, if teachers know what is expected of them, they will do it.” While Teacher 9’s statement seems intuitive, the participants’ responses indicated that the theoretical simplicity of communicating clearly defined expectations is often not as simple in practice.

Teacher 10 reflected upon the specificity of the model and its impact on both lesson planning and administrators’ expectations: “I think it’s streamlined my teaching. It’s also streamlined what they are looking for.” However, Teacher 10 did indicate that the narrow focus has the potential to negatively impact a teacher’s evaluation. She stated the following:

They’re much more focused on two or three elements. I think there are pluses and minuses to that. I think if they come in trying to see three particular elements, and they don’t see those elements, I think that could be detrimental to the teacher.

Teacher 14 expressed the same concerns, stating, “I don’t think [the Marzano Model] covers enough. It’s narrow.” In their responses, both Teacher 10 and Teacher 14 spoke to the issue of teacher autonomy. Teacher 5 also indicated that the model helps “set clear parameters” to aid in establishing clarity of administrators’ expectations. Four of the eight teachers that mentioned clarity of expectations in their responses communicated positive opinions about the specificity of the Marzano framework; however, every one of these four teachers spoke to the limits of the model as well, alluding to a connection between clear expectations and lack of teacher autonomy.

While the Marzano Model is intended to provide a framework from which teachers can plan, implement, and reflect upon instruction, four teachers perceived there was a lack in clarity
of district expectations after the model was implemented. Teacher 7 expressed concern about the subjectivity of the evaluator, saying, “I don’t know if it’s just here, but you always have this fear like they could come in and observe whatever the heck they want on Marzano.” Five of the 14 participants expressed that the Marzano Model was “overwhelming” due to the incorporation of 41 observable elements. Teacher 6 echoed the sentiments of Teacher 7 as well as the comprehensive nature of the Marzano Model:

I guess the only issue that I’ve had with it is just when I was being observed, sometimes it’s kind of hard to know which segments the observer wants to see, and I guess my concern always is at any given time you are looking at these 41 elements, and you would have to be having a pretty awesome day to hit all of these in one lesson.

While Marzano (2011) asserts that evaluators may not, and in fact should not, be observing all 41 elements in a lesson, the perceptions of the participants in this study reflect that the sheer number of possible elements continues to overwhelm teachers.

**Depth and Breadth of the Evaluation**

With regard to the Marzano Model’s accurate depiction of professional performance, depth and breadth of the district’s application of the instrument, specifically relating to administrator’s knowledge of teacher capabilities, remained a concern for the participants. Eight of the 14 participants voiced their concerns that the observations of their evaluators did not reflect an accurate picture of their professional capabilities due to the fact that these observations represented a minute portion of a very large puzzle. Teacher 10 provided a very poignant observation that reflected her concerns regarding her perception of the accurate depiction of professional capabilities: “I feel like because someone comes in here twice a year and sees one thing, they just don’t understand [what I’m doing in my classroom]. I feel like people here don’t
even know me. I feel judged sometimes, but no one knows me.” This comment speaks to the teacher’s perceived lack of depth in the district’s use of the Marzano Model. Ultimately, teacher evaluation is utilized in an attempt to “know” the teachers’ strengths as well as weaknesses and identify strategies to assist in professional improvement and growth.

Voicing similar concern for accurate reflection of professional capabilities, Teacher 9 said, “It’s tough when there are people who are doing everything in every way, and I’m not saying I am, but some people do more than is expected but are not seeing the reflection in their observations.”

Based on the responses of the participants, the breadth of the tool, or the ways in which the Marzano Model effectively captures teachers’ professional performance over the course of the year, is directly related to the number of observations performed throughout the year. Six of the participants mentioned the number of observations performed throughout the school year and the relative connection to the accuracy of the professional evaluation. Teacher 4 suggested, “More would be better because you just get an idea of time to time. Three snapshots are better than one, but it could be that you came in on the three weirdest lessons I did that year. You can’t help it, it just happens.” Teacher 4 indicated that he felt that administrators were “spread too thin,” but expressed the importance of frequent visits from administrators.

In a tone that reflected dejection and disappointment, Teacher 12 lamented:

Sometimes I have a problem because it’s hard when you evaluate me in one class, but you don’t see what I do everyday. And these evaluations come back sometimes and you’re like, “Well, how do you know that I’m not doing that already?”

The concerns expressed by Teacher 12 surfaced in four other participant responses as well.
Constructive and Useful Feedback

Participant responses indicated that specificity was the one major theme that emerged relating to effective feedback that fosters and promotes professional growth under the Marzano Model. Districts that adopt the Marzano Causal Teacher Evaluation Model utilize the iObservation platform to record, organize, and manage data collected from classroom observations. Within iObservation, the evaluator has the ability to “check off” specific boxes that indicate the occurrence of specific, observable behaviors (by either the teacher, the students, or both). These observable items or occurrences are codified as “evidence” within iObservation.

Seven of the 14 respondents perceived that the Marzano Model had strengthened evaluator feedback as a direct result of the documentation of specific evidence observed during a lesson. Teacher 6 stated, “I think we are getting more helpful feedback because we do have those [evidence] items on our observations under the elements; whereas before, there weren’t necessarily those specific points for improvement.” These seven participants viewed the incorporation of specific evidence as a clearly defined list of teacher and student behaviors that the existing research proves are linked to an increase in student achievement.

In an effort to streamline the district administration’s approach to the 41 observable elements and their accompanying pieces of evidence, district administration communicated that evaluators would limit the number of elements selected to between two and five, depending upon the length of the observation. However, Teacher 1 expressed one downfall to the specificity and deliberate limitation of the feedback provided from the use of the Marzano Model: “When you have six different observers, I guess our hope was that out of those 41 elements, they would see more than just the same three all of the time.” Specificity in feedback is essential when
attempting to directly inform teacher instructional practices; however, if the focus is too narrow, it is possible that the limitations could, in turn, limit professional growth.

Perceptions of Proficiency Scales and Ratings

Perhaps the most consistent response among the participants was the identification of scales and scores as barriers to professional growth. Thirteen of the 14 participants discussed the Marzano proficiency scale with respect to teacher professional practice as an implement that hinders professional growth as opposed to fostering the desire to improve. Twelve of these 13 teachers discussed the impact of the communicated difference between Applying and Innovating on the Marzano proficiency scale.

In order to provide an accurate report of the data collected, it is imperative to understand the way in which the proficiency scale is categorized. The Marzano Causal Teacher Evaluation Model provides a 5-point scale for each of the observable elements that includes the following levels: Not Using (0), Beginning (1), Developing (2), Applying (3) and Innovating (4). As an example, one element that is listed in Domain 1 of the Marzano Model is “Processing New Information.” The proficiency scale indicates that a teacher who demonstrates behavior of Applying “engages students in summarizing, predicting, and questioning activities and monitors for evidence of the extent to which the activities enhance the majority of students’ understanding.” The Innovating teacher “adapts and creates new strategies for unique student needs and situations in order for the desired effect to be evident in all students.”

When asked about the accuracy of the proficiency scale with respect to professional capabilities or performance, all of the 14 teachers discussed the existence and/or attainability of the Innovating rating. Of the 13 teachers that specifically discussed Applying and Innovating ratings within the Marzano Proficiency Scale, six suggested that the Innovating level is largely
unattainable in their district, five noted a lack of clarity as the primary source of frustration, four identified a dip in morale as a result of the district’s use of the scale, and one teacher suggested that the proficiency scale limits professional growth. Four teachers mentioned both lack of clarity and the perceived unattainability of the Innovating score in the district. Two teachers mentioned both the unattainability of the Innovating score in the district and teacher morale within the district.

Teacher 8 contrasted the administrator’s approach to the Marzano Model to her approach to her own classroom:

It’s hard to be objective because we were basically told that they will never give out Innovatings. Now, I know that’s changed a little bit; it has softened a bit. When they first started four years ago, it was [communicated that] Innovating is unattainable. It was there for something for you to reach, and to me, that was just . . . as a teacher, I would never tell my student, here’s the 100, here’s the A+, but you’re never ever going to see it. She spoke to the perceived unattainability of receiving the Innovating rating for an observed element.

When asked about her perception of the proficiency scale rating, Teacher 12 touched upon lack of clarity with administrator expectations as well as the perceived unattainability of the Innovating score. Teacher 12’s response, as well as the tone associated with the response, reflected her genuine misgivings with the proficiency scale:

I don’t agree with it, and it does make me feel bad sometimes. Sometimes, I’m like, “Why don’t I ever get Innovating?” And then [my supervisor] gives me examples, and I’m like, “But I’ve done that, and what happened?” It’s hard when you’re being
evaluated by so many different people and they are looking at you, and there are so many different ways of looking at what a good and effective teacher looks like.

More than half of the participants posed the question: What does Innovating look like? When asked a follow-up question to define the district’s expectations with respect to the Innovating rating, three of the eight teachers to whom the question was posed answered by suggesting that they felt that they must do something new and different in order to be rated as Innovating on the proficiency scale. The remaining five teachers expressed concern that the district definition of Innovating was vague, at best.

With respect to morale, Teacher 8 discussed the impact that the proficiency scale has had on district morale:

If you are never getting an “A” and you’re always getting “B’s,” as a teacher, I just think that it’s poor practice to have an evaluative tool where you can’t get an “A.” And I don’t like that at all on evaluations. I just don’t think it’s a good way to motivate people to do better because it says that as long as you’re okay, you’re okay, but you’re never going to get up there.

Eight of the 14 teachers associated the proficiency scale with the grading scale that they use to assess their students. Teacher 3 reflected upon her perception of the district’s expectations regarding the proficiency scale as well as the definitions associated with the terms within the scale:

I think maybe the definitions of those could be looked at, and if it can’t be looked at, then maybe it could be explained to us that Innovating is not an “A,” Applying is not a “B,” Developing is not a “C,” and Beginning is not “you’re out the door.” Perhaps it could be defined a bit differently for us.
Teacher 1 provided a generalization of teachers in her assessment of the proficiency scale. She remarked, “We, as teachers, we were the go-getters as kids. There is a reason we chose to do what we are doing. I loved being in school. So when someone says, “You’re a 91,” you know, we cringe. “Well, why aren’t I a 92?” So, yes, I think that psychologically, it represents a barrier.”

Three other teachers in addition to Teacher 1 shared a similar opinion of teachers. These participants believed that most teachers were excellent students, which propelled their desire to enter the field of education. Teacher 2, for example, said that teachers “want to do everything right” with respect to instructional practices. According to Teacher 2, if teachers are not constantly rated as Innovating, they feel as though they have failed.

When asked the same question, but with respect to the summative evaluation rating (Highly Effective, Effective, Partially Effective, and Ineffective), teachers had a similar mindset. Only one of the 14 participants was rated Highly Effective in the summative evaluation. All of the other 13 teachers were rated Effective. Five of the 14 participants perceived the Highly Effective rating to be unattainable. Teacher 11 reflected upon the impact the Highly Effective rating has had on the teachers in the district. In a tone wrought with lament, she said, “I think that the big problem with Marzano is the fact that nobody can get Highly Effective. I think that when you tell a faculty that there is nobody here that is Highly Effective, I think that’s what’s hurting [the morale].”

Teacher 13 had a distinctly different tone when addressing her perceived issues with the district’s low number of teachers that are rated Highly Effective. She joked, “What do I have to do, jump through hoops of fire [to be Highly Effective]?” While teachers discussed their distaste of the summative evaluation rating scale in varying degrees, the majority of the teachers
remarked that the summative evaluation ratings have had a negative impact on their experiences as teachers in the district. These teachers reported a lack of willingness to take risks, a feeling of being indistinguishable amongst their peers, and an overall feeling of being inadequate.

**Summary of Findings for Research Question 3**

The present study speaks to the fact that the relationships between teachers and administrators are complex and can be influenced by any number of factors. Teachers at Silver Creek reported largely positive relationships with their administrators. However, the way in which the change to the Marzano Model was implemented greatly impacted the relationship between teachers and administrators, instilling a sense of fear and mistrust among the teachers. Participant responses also reflected the fact that broad-based change should be implemented with great care. The change to the Marzano Model negatively impacted the relationships within the organization. Despite the fact that the administration responded to the criticism of the implementation of the model, teacher responses reveal that it is difficult to remedy negative associations with the change after it has been implemented.

Participant responses also implied the importance of building trust between teachers and administrators through demonstrated competency among the administrators. It is less likely that administrative feedback throughout the evaluation process will have an impact on teacher growth if teachers perceive administrators to be inept or lacking experience in the content.

Finally, the present study reflected the value in viewing administrators as coaches over evaluators. Teachers valued informal feedback just as much, if not more, than formal feedback provided in post-conferences. If teachers perceive the informal observations as minimally risky, they are more apt to take instructional risks, and ultimately, demonstrate true professional growth.
**Themes: Research Question 3**

Research Question 3: What are teachers’ perceptions of the influence of the Marzano Causal Teacher Evaluation model on the relationships with their administrators?

Interview Questions 8-10 were designed to address Research Question 3 through discovering teachers’ perceptions of their relationships with administrators and how the district’s implementation of the Marzano Model affected these relationships. In general, eight teachers felt that they had a positive relationship with their administrators, two felt that they had a negative relationship with their administrators, and four were neutral in their assessment of their relationship. Themes that emerged as a result of asking the interview questions related to Research Question 3 were coaching versus evaluating, trust, and understanding and managing change.

**Coaching Versus Evaluating**

Six teachers identified trust as an essential component to establish and maintain rapport with administrators. Every one of the eight teachers who reported that they had a positive relationship with their administrators believed that trust was built over time. At the time that this study was conducted, the district was piloting Project COACH, an evaluation model that linked teachers to one evaluator throughout the year. The teacher’s evaluator performed 6-8 shorter (10 minute) observations throughout the school year but utilized the proficiency scale to rate the teacher only twice. The Marzano Model was still utilized throughout the process, but the evaluators acted as coaches throughout the year, offering teachers targeted feedback and regularly checking growth over a period of time. Teachers were evaluated once at the midpoint check-in and once at the end of the school year.
Participation in the pilot was strictly voluntary. Four of the 14 participants in this study volunteered to be part of Project COACH. Three communicated a preference for the COACH model over the district’s traditional evaluation model, while one teacher preferred the traditional approach.

Teacher 1, a veteran teacher of over 20 years, expressed her excitement for the altered approach to teacher evaluation: “I was excited about Project COACH because I thought that if there was something that my observer saw in the first lesson, that’s something that she and I could work on for the rest of the year.” Teacher 1’s response speaks to the need for more frequent observations as well as a more targeted and individualized approach to teacher evaluation.

Teacher 11, another participant in Project COACH, spoke specifically about the improved relationship between herself and her evaluator under the COACH model. “I think the relationships that are established in the coach model, the administrator has a better understanding of what the teacher is doing, so you have better conversations about your teaching, and better conversations about instruction through that.”

This comment is a stark contrast to Teacher 7’s comment that she felt “judged” but felt that no one truly knew her. Interestingly, Teacher 7 participated in the COACH model for the 2015-2016 school year but indicated that she had opted out of the program for the 2016-2017 school year, citing the need to regain a sense of normalcy.

During the 2015-2016 school year, the Silver Creek School District also made efforts to assign all tenured teachers to one evaluator as opposed to obtaining multiple perspectives from multiple evaluators throughout the year. This approach, while not identical to Project COACH, reflected an approach that was more akin to coaching than evaluating. Four participants revealed
that they preferred one evaluator over multiple evaluators throughout the course of the school year. Teacher 13 said, “This was the first year when you had the same evaluator three times, which was great because it felt like you grew with the evaluator.” She went on to reference a particular experience with her evaluator during which the evaluator passed along advice but did not formally document the observation or the conversation. Teacher 13 felt the single evaluator had a better chance of getting to know her as not only a teacher, but as a person who is very invested in her students and is continually seeking professional growth.

Teacher 4 echoed a similar sentiment regarding a preference for a single evaluator. “I actually liked having the same person a little bit better because then you can kind of talk about, “I took your advice here; is there anything I can do to make that better?” Teacher four’s comment speaks to the need for follow-up, low-stakes conversations between teachers and evaluators.

Trust

The participants were largely positive about their relationships with the administrators within the district. Often, teachers identified with the presumed struggles of evaluators to effectively and accurately evaluate teachers. Seven of the eight teachers that felt that they had developed positive relationships with their administrators since Silver Creek’s implementation of the Marzano Model reported that they valued the professional conversations that took place between themselves and their evaluators. It is human nature to value feedback from an individual that you trust. Teacher 3 drew a connection between professional conversations and trust. She stated, “[The structure of the model] helps us to share a lot more, and it helps us to build more trust between us and administrators, which is all really important. That feedback has the opportunity to bring us really close.”
Teacher 8 also spoke to the importance of trust:

That conversational piece is so important, but it kind of works to your advantage if you know the person, as it would in any situation. My experience has been nothing but positive with every administrator that has observed me because I’ve known all of them and trusted all of them.

Every single one of the 14 teachers interviewed indicated that they valued the feedback they received throughout the evaluation process. Even when teachers were hesitant to fully support the district’s decision to implement the Marzano Model, they expressed the fact that they have taken the feedback they have received throughout the evaluation process and applied this feedback to their classrooms. Due in large part to the specificity of the model, the majority of the teachers believed that the feedback from their administrators had improved since the implementation of the Marzano Model. What’s more, the perceived value in this feedback has helped to strengthen professional conversations and, therefore, professional relationships between teachers and administrators.

Only one teacher preferred the “old” model the district had utilized prior to the implementation of Marzano. Teacher 14 detailed the parameters of the former model as well as the ways in which this model was more beneficial to her professional growth: “[The Marzano Model] is narrow. I mean, we used to have a system here . . . it was a big, long checklist. It probably was a pain for the administrators. In each section, they would put comments. That was extremely helpful to me.” The model referenced by Teacher 14 was a comprehensive checklist that administrators would utilize to “check off” Satisfactory, Unsatisfactory, or Not Using in reference to a large number of observable behaviors within the classroom. Beneath each section of behaviors, administrators would record comments that expounded upon the
observable behaviors. The Marzano Model is not much different, with the exception of the proficiency scale and the idea that not all of the 41 elements will be observed in one particular lesson.

One final component of trust that surfaced during this study was the teachers’ willingness (or lack thereof) to accept the administrators as experts, thus influencing the trust teachers had in the decisions their administrators were making as well as the applicability of the feedback they were providing. Four of the participants referenced the importance of the experience the evaluator has had in the classroom, specific to the content areas of the teachers they are evaluating. Teacher 4 remarked as follows:

Also, the background [matters]. I mean, some people have been doing special education their whole lives, so they understand that perspective, and some people . . . it’s always going to be hard for someone who doesn’t have a math background to come in and say, “You taught that math concept wrong.” They are going to have to focus on other areas of my instruction.

Teacher 14 agreed with this assessment. Upon reflecting on the emphasis on instructional practices over content, she said, “And then people say, well it’s about strategies and practices, but not really. That’s where the science and art thing is . . . each content area is different. I don’t know. The business model and education don’t go hand in hand.” Teacher 14 alluded to the fact that good performance does not necessarily look the same across content areas. If teachers do not trust that administrators are competent to evaluate them due to their lack of experience in the content, it is less likely that teachers will value the ever-important feedback they receive throughout the evaluation process. Trust between administrators and teachers, therefore, is multi-faceted, but boils down to perception of competency and relationships.
Understanding and Managing Change

When asked to reflect upon their experiences with their administrators throughout the process of adopting and implementing the Marzano Causal Teacher Evaluation Model at Silver Creek High School, the one theme that emerged that reflected a negative experience for teachers was understanding and managing change. According to most of the teachers interviewed, the district implemented the model at a pace that did not reflect a willingness of district administrators to accept the fact that teachers would not be instantly comfortable with the Marzano Model and its prescribed instructional practices. Seven of the 14 participants stressed the importance of allowing teachers time to interact and experiment with a teacher evaluation model prior to strict implementation.

Upon reflection of her opinion towards the Marzano Model, Teacher 8 responded as follows:

I think part of the distaste I have for Marzano is not necessarily the model, it’s more of the way it was implemented at Silver Creek. There has been so much change over the past three years; I still feel sometimes that I can’t keep my head above water. While I understand it was required by the state that we choose one of these models, I think the implementation was done too quickly. The stress and the pressure that was put on . . . there wasn’t a whole lot of wiggle room that first year. So I think teachers were just scared. You had teachers that had been teaching for 20 years, and now there’s this change, and this change was not necessarily done gently.

Teacher 8 was not alone in her criticism. Teacher 7 expressed her belief that the administration was “hypercritical” from the onset of the implementation of the Marzano Model.
She also indicated that many teachers felt that they were being “dinged” by administrators on the evaluation model before they felt comfortable with the district’s expectations.

Four teachers mentioned the district’s approach to training as a major flaw in the implementation of the change to the Marzano Model of teacher evaluation. Teacher 10 explained her perception on the initial Marzano training:

We had two big district meetings. With a hundred and some odd teachers. And it was really awful. Here we were talking about these elements, and we have no clue what they are. I was like, “Okay, this meeting is not helping us.” I felt like it should have been more about the elements and the teaching style. So I think if you ask any teacher in Silver Creek, we were never properly trained.

Teacher 10 also indicated that she felt that the administration was not properly trained to utilize the evaluation model, either. Teacher 1 disagreed, indicating that the administration was doing “as much as they could” to learn the Marzano Model and make deliberate efforts to calibrate their interpretations of the model through dual observations.

Finally, teachers referenced the fear associated with the change to the Marzano Model. Seven of the 14 participants indicated that they felt stressed and overwhelmed at the onset of the district’s implementation of the Marzano Model. The majority of these seven participants reported that there is still a feeling of stress and anxiety towards teacher evaluation. Teacher 12 revealed that she feels that she is “going through the interview process all over again” during both pre- and post-observation conferences. Teacher 3 said that she is always “a nervous wreck” when an administrator is in the room. She also addressed her hesitancy to take risks in the classroom due to the current teacher evaluation system in place. This particular participant had many positive things to say about the Marzano Model, her administrators, and teaching in
general; but she did allude to the fact that she is less inclined to try something new for fear of a poor evaluation: “We are less prone to take a risk. We are also putting ourselves out there. You fear losing your job or you fear just not looking as good as how hard you have worked.”

The most poignant statements regarding the stress and anxiety associated with teacher evaluation at Silver Creek were addressing the district’s first year of implementation of the Marzano Model. Teacher 4 provided insight as to the overwhelming nature of the change. He said, “I think it is because it was so different from what we were doing before, and there is so much to it that it was a little overwhelming.” The Marzano Model represented a paradigm shift in this district, much as teacher evaluation reform did across the country. It was difficult for teachers to immediately adjust to the change, thus creating an environment of fear and, ultimately, mistrust towards the district’s administration.

Teacher 7 attributed teacher stress to the scales in particular. Again, she alluded to the district’s initial implementation of the Marzano Model: “I don’t mind that there is a scale; it’s really the way that it was implemented. It was kind of unfolded in an abrasive way where it made people very nervous and stressed out.” Teacher 7 expounded upon this statement, recalling teachers that have “cried and fought” over getting Developing scores in their evaluations.

However, the remainder of the participants that discussed the amount of stress associated with teacher evaluation indicated that the administration did reflect upon the impact the implementation of the Marzano Model had on the teaching staff and adjusted accordingly. For example, three of the teachers referenced the “binders” that were required in the first year the district utilized the Marzano Model. Teacher 7 recalled the time-consuming task of collecting and “proving” all of the contributions he had made to the district through filling a binder full of
“artifacts.” The district administration did not require teachers to submit the binders after the first year due in large part to the feedback they received from the teachers.

There was another distinct shift in focus and change in requirements between years two and three of the district’s implementation of the Marzano Model. During Years 1 and 2, teachers were evaluated using all four domains of the Marzano Model. However, beginning in Year 3, the district administration shifted to exclusively evaluating teachers using the 41 elements in Domain 1 of the model. The administration claimed that all of the other three domains would be directly reflected in a teacher’s instructional practice. These changes did help to ease the stress of the teachers. However, the relationships between teachers and administrators did suffer as a result.

Summary

Chapter IV presented an overview of the findings of the research study through answering each of the three research questions in an attempt to answer the overarching research question: What are teachers’ perceptions of the Marzano Causal Teacher Evaluation Model? Chapter V presents an in-depth discussion of these findings with respect to the theoretical frameworks that guided the study. Additionally, Chapter V addresses considerations for policy and practice.
CHAPTER V
CONCLUSIONS AND RECOMMENDATIONS

Introduction

This chapter presents an overview of the present study. It begins with a summary of the purpose of the study as well as a restatement of the research questions. This is followed by a summary and discussion of the study with respect to the review of the existing literature and the theoretical frameworks from which the study was structured. The chapter concludes with future considerations and recommendations for practice, policy, and research.

The purpose of this study was to explore how teachers perceive the influence of the Marzano Causal Teacher Evaluation Model on their professional practices. The research indicates that teacher effectiveness is one of the most significant factors in student achievement (Garrett & Steinberg, 2015; Heck, 2009; Sanders et al., 1997; Stronge et al. 2007). When addressing teacher quality, one critical difference between ineffective and effective teachers is the fact that effective teachers are reflective practitioners, continually seeking to develop and cultivate highly effective instructional practices (Black & Howard-Jones, 2000; Stronge, 2007). The purpose of teacher evaluation, therefore, is to indirectly impact student achievement through directly impacting teacher performance. Through developing an understanding of the ways in which teachers perceive the evaluation tool and process, administrators should be able to better understand and implement effective practices when conducting evaluations.

Teacher perception of evaluation is dependent upon the perceived credibility of the evaluator as well as the evaluation tool. Employees are more likely to accept the feedback provided by an evaluator as accurate and make the suggested changes if the employee believes the evaluator has successfully implemented the evaluation tool with fidelity (Kinicki et al.,
If administrators are aware of teachers’ perceptions relating to evaluation, they could use this information to inform their policies and protocols associated with teacher evaluation, thus positively impacting teacher performance and, ultimately, student achievement. It is imperative, therefore, that school districts consider teacher perception of evaluation when implementing and maintaining a teacher evaluation tool.

The sample for this qualitative study was 14 high school teachers at Silver Creek High School. I conducted one-on-one semi-structured interviews with each of the participants. The research sample consisted of two males and 12 females with four to 23 years of teaching experience. The participants represented a variety of content areas, including World Language, Math, English, Social Studies, Visual and Performing Arts, and Science. Two participants were certified Special Education teachers. All 14 teachers were tenured. This research was conducted during the summer of 2016.

**Research Questions**

The following three questions were used to guide the phenomenological qualitative research design:

1. How, if at all, has the Marzano Causal Teacher Evaluation Model influenced and informed teachers’ instructional practices?
2. What are teachers’ perceptions of how accurately the Marzano Causal Teacher Evaluation Model reflects and captures their professional performance and capabilities?
3. What are teachers’ perceptions of the influence of the Marzano Causal Teacher Evaluation Model on the relationships with their administrators?
Summary and Discussion of Major Findings

The changes and adjustments in teacher evaluation in New Jersey since the adoption of AchieveNJ in 2013 have elicited attention from teachers and administrators alike. In an attempt to increase teacher accountability and, in turn, student performance, policies requiring school districts to implement high-stakes teacher evaluation were implemented across the country. The teachers interviewed for this study perceived that an evaluation model with a clearly defined framework, like the Marzano Causal Teacher Evaluation Model, could aid in improving teacher performance and increasing student achievement. The manner in which the evaluation instrument is implemented is indicative of the future success of the model. This study has the potential to shape the policies and practices related to teacher evaluation through capturing the human component of current evaluation practices in a large New Jersey high school.

Findings Related to Research Question 1

Research Question 1 states, How, if at all, has the Marzano Causal Teacher Evaluation model influenced and informed teachers’ instructional practices? The literature suggests that the quality of instruction, more specifically the quality of the teacher planning and implementing instruction, has a direct impact on student achievement (Sanders, Wright, & Horn, 1997; Sanders & Horn, 1998; Darling-Hammond, 2000; Heck, 2009). It is imperative, therefore, to consider how teachers perceive the evaluation model and its influence in affecting teacher practice.

With respect to lesson planning, the participants’ responses indicated that the Marzano Model has added structure to their lesson plans. A majority of the teachers indicated that the Design Questions (DQs) in the Marzano Model influenced their approach to lesson planning in a positive way. Most of the teachers felt that they were more deliberate in their planning and more
reflective in their practice as a result of the structure and design components (DQs and 
is instructional strategies) of the Marzano Model.

However, more than half of the participants indicated that the lesson plans they submit 
are solely for administration, suggesting that compliance was a significant motivation in lesson 
planning. When the district began using the Marzano Model as the formal evaluation tool, the 
lesson plan requirements changed. Teachers were required to include the DQs, specific 
elements, and learning goals in their lesson plans. While teachers appreciated the structure this 
model provided, they did not feel as though their instructional approach had significantly 
changed.

The responses of the participants also revealed teachers’ desire to be reflective in their 
practice and improve as teachers. The Marzano Model not only encourages but also expects 
teachers to incorporate reflection into their practice. All of the 14 participants expressed their 
desire to be quality educators. Teachers were very open about their desire to grow, citing the 
fact that they were always willing to make changes in their practice because they by no means 
were “perfect.” The majority of the teachers indicated that their teaching has improved since the 
implementation of the Marzano Model due to the model’s emphasis on reflective practice and 
student achievement.

Similarly, all of their responses connected student success and achievement to the quality 
of their instruction. As previously noted, Marzano (2011) cites student achievement as the single 
most important factor in assessing teacher quality. Teacher evaluation instruments have 
reflected the value in accounting for student achievement for almost a century (Wetzel, 1929). 
All of the teachers reported that they checked for student understanding more frequently, while 
most of them believed that they were more focused on how they were getting their students to
achieve the learning goals. Most teachers also reported that the changes in the curriculum from fact-based to skills-based since the district’s adoption of the Marzano Model were influential in making gains in student achievement. These findings speak to the importance of deliberate and effective instructional design. Marzano’s framework provided teachers and administrators with a clearly defined framework from which they could develop goal-oriented curricula.

Teachers valued the feedback from their administrators, indicating that this feedback was more specific and detailed than in years prior to the implementation of the Marzano Model. Marzano (2003a) lists feedback to teachers as one of the five school-level factors that influence student achievement. Teachers must be informed of what they are doing in the classroom and how they can further tailor their instruction to meet the needs of their students. Formative assessment is a component of all state-approved evaluation models due to the positive impact it can have on student achievement as well as teacher practice. Teachers reported an increase in their use of formative assessment and supported the fact that formative assessment has been useful in guiding their instruction.

One additional criticism that emerged from the analysis of the participants’ responses was the fact that they felt that this model negatively influenced teacher autonomy, prompting administrators to treat all teachers in the same way, regardless of the teachers’ content area or teaching style. While teachers expressed that they appreciated the evidence Marzano lists for observable teacher practices, their responses also reflected a need for a more flexible approach to teacher evaluation that accounts for the nuances of teaching and learning. Several teachers equated their experience with the Marzano Model with being “a square peg in a round hole”; teachers felt that they did not always fit the model, but they persisted in making changes to their
instruction that they deemed unnecessary and, at times, ineffective in an effort to meet the perceived expectations of the administrators in their district.

Findings Related to Research Question 2

Research Question 2 states, What are teachers’ perceptions of how accurately the Marzano Causal Teacher Evaluation model reflects and captures their professional performance and capabilities? Both the literature and existing policies reflect the following two purposes for teacher evaluation: (1) to improve teacher instructional practice, and (2) to inform personnel decisions (retention and dismissal) (Donaldson & Papay, 2014; Gabriel, 2015; Garrett & Steinberg, 2015; Marzano & Toth, 2013). The Marzano Model was implemented in Silver Creek High School in order for the district to maintain compliance with state mandates. As a result of the changes in teacher evaluation policy in the state of New Jersey, the Marzano proficiency scales and summative evaluation ratings were more than just hierarchical labels; they were a representation of job security for teachers. For the first time in the existence of tenure, it was possible for tenured teachers to lose their tenure status and, ultimately, their jobs. It is not surprising that the teachers that participated in this study expressed negative feelings about the proficiency scale and summative evaluation.

During the year that the study was conducted, 80% of teachers’ evaluations were comprised of teacher practice, while the SGO score accounted for 20% of the teachers’ summative evaluation score, reflecting a clear emphasis on instructional practices. With respect to the proficiency scale, teachers were not opposed to Marzano’s inclusion of a proficiency scale in the evaluation model. Their responses reflected quite the opposite. They appreciated the scale and its accompanying definitions. They agreed with the need for a scale to communicate and enforce professional standards as well as maintain a system of teacher accountability.
For example, one teacher who had been teaching for 16 years said that he never had a problem with accepting constructive criticism that was warranted. He provided an example of a lesson in which he was rated as Developing in one particular element. He agreed with his administrator’s assessment. Some teachers perceive accountability as a necessity in the education profession. The current study intimated that teachers often value equity and fairness over praise and accolades with respect to teacher evaluation.

The majority of the teachers had an issue with the administrators’ interpretation of the proficiency scale, many indicating that they were told that administrators were going to “give out” Innovating ratings sparingly. Teachers equated the scale with a grading system, suggesting that Innovating is an “A,” Applying is a “B,” Developing is a “C,” Beginning is a “D,” and Not Using, is an “F.” The large majority of the teachers were rated as Applying and Developing, fostering the perception that they are not “A” students, not high quality educators, but rather, mediocre. The participants associated the implementation of the Marzano Model to a dip in morale because teachers felt that the scores they were receiving, both during observations and as their summative evaluation scores, were not reflective of the amount of work they were putting into teaching. They felt that the district’s implementation of the Marzano Model had created a more regimented and cumbersome approach to instruction, thus prompting an increased workload, but a decrease in their administrators’ perceptions of their abilities and expertise.

Participant responses also indicated a need for increased inter-rater reliability with respect to the proficiency scale ratings. While the AchieveNJ Act requires administrators (teacher evaluators) to receive an initial training and subsequent “refresher” trainings on the district-selected evaluation instrument in addition to a minimum of two dual observations in
order to promote inter-rater reliability, the fact remains that teacher evaluation is subjective and, therefore, somewhat lacking in reliability (Herlihy et al. 2014; Peterson, 1987, 2000).

The majority of the teachers reflected upon the subjectivity of teacher evaluation, indicating that it is a regular struggle to determine what the evaluator is “looking for” when they are observing a classroom as well as how they apply the ratings. One teacher provided an example of a lesson that was observed during two separate years by two different evaluators. She received a score of Applying for one particular element from one administrator and Developing for the same element from a different administrator during the subsequent school year. She felt as though her lesson plan had been “downgraded” and reflected upon why she does not have complete trust in her administrators to fairly and equitably utilize the evaluation model. Instances such as this speak to the importance of regular calibration among administrators, a practice that is often overlooked in a sea of policies, protocols, and regulations.

With regard to the summative evaluation rating, teacher responses primarily mirrored those relating to the proficiency scale: the majority of the teachers felt that their summative evaluation was not reflective of the amount of effort they put into teaching. During the first year of implementation, teachers reported that only one teacher in the entire Silver Creek School District, which staffs close to 200 teachers, was rated Highly Effective. I do not have absolute confirmation that this is the case, but according to the New Jersey teacher ratings report released in 2013-2014, 102 of the 107 teachers working at the high school during the 2013-2014 school year were rated as Effective. The other ratings (Ineffective, Partially Effective, and Highly Effective) were not disclosed; however, only five remaining teachers could fall into one of the unpublishized categories.
Teachers that participated in this study expressed their concerns that they feel that they are all lumped together in one big Effective group. Many teachers expressed their desire to be Highly Effective, describing this rating as the “A” they are striving for, much like Innovating on the proficiency scale. Participants’ perceived the summative evaluation rating as a barrier to their professional capabilities and growth. Teacher belief of job capabilities, or teacher efficacy, is related to teachers’ enthusiasm, persistence and perseverance, commitment to teaching, and instructional behaviors (Tschannen-Moran & Hoy, 2001). If teachers believe that they are less capable of performing their jobs, they will be less likely to commit to adjust and improve instructional behaviors, thus negating the purpose of teacher evaluation. Interestingly, not one teacher I interviewed expressed discontent with the Marzano Model exclusively. Teachers were mostly complimentary of the model, explaining that they believe the structure the model provides has been largely beneficial for teachers in the district.

**Findings Related to Research Question 3**

Research Question 3 stated, What are teachers’ perceptions of the influence of the Marzano Causal Teacher Evaluation model on the relationships with their administrators? The majority of the participants valued the professional conversations that took place after the district implemented the Marzano Model. The specificity of the model guided the professional conversations between teachers and administrators. The evaluation model provided administrators and teachers with a common language from which to assess and reflect upon teacher practice. While some participants identified the instrument’s specificity as a negative attribute of the model, the majority of the participants praised the framework for providing them with specific research-based instructional strategies to incorporate in their classrooms. Most of the participants believed that the Marzano Model had improved their teaching skills. The
specific framework of the Marzano Model helped to improve the quality of professional conversations, thus improving the professional relationships between teachers and administrators.

Teacher responses also revealed that trust is a primary factor in cultivating teacher-administrator relationships. The same is true for any human relationship. While the majority of the participants reported that they had good relationships with their administrators, their responses reflected a belief that the trust between administrators and teachers was negatively influenced by the implementation of the Marzano Model. Much of this can be attributed to the way in which the district implemented the model during the first year. Teachers reported feeling “stressed” and “overwhelmed” to the point where some participants reported the health problems that they and a few unnamed colleagues experienced during the district’s first year of using the Marzano Model to evaluate teachers. Teachers felt inept and ill equipped to effectively instruct their students while being evaluated under the Marzano Model, due in large part to the training they received prior to the implementation. They felt as though they were being evaluated under a system with which they were not yet comfortable and, at the same time, fighting to keep their positions within the district. This fear permeated the district, and unfortunately, the administrators were blamed for the majority of the flaws in the newly adopted evaluation process.

Of course, the Marzano Model has become synonymous with the teacher evaluation ratings and NJAchieve due to the fact that the model was used as a means to evaluate teachers under to new state-mandated evaluation requirements. However, teacher perceptions of the first year of the model were and continue to be largely negative. Several of the participants perceived that administrators were going out of their way to criticize teachers instead of encouraging their
professional growth. These participants also felt that each administrator was “looking for” something different, a belief that negatively impacted the trust between teachers and administrators. In these ways, the model thwarted administrators’ efforts to continue to foster positive relationships between the teachers and themselves.

A final significant finding involved the distinction between the district’s approach to a modified teacher evaluation process: Project COACH. All but one of the teachers that participated in Project COACH indicated that they felt that their relationships with their evaluators were strong. Teachers reported that the COACH program fostered a more trusting relationship between the teacher and the evaluator. These teachers viewed their assigned evaluators as a support system rather than a judge. It is significant that the same Marzano Model was used to evaluate these teachers throughout the process. The only difference between the two models is that the COACH Model incorporated a higher frequency of shorter observations coupled with fewer formal evaluations.

**Findings Related to the Theoretical Frameworks**

**Adult Learning Theory**

Adults learn differently than do children due to the fact that adults are often “relearning” the content being presented. Adults are often required to reconcile previous beliefs with newly constructed beliefs. This study is a prime example of “relearning” content. Teachers were once exposed to the former evaluation model, a model that had stood the test of time. For decades, school districts utilized the evaluation model that rated teachers’ observable behaviors as Satisfactory, Unsatisfactory, or Not Using. When the Marzano Model was introduced, teachers had to construct a new working definition of teacher evaluation. They built their definition, as adults do, around their experiences with the model, both independently and as a faculty.
Teachers were now being assigned an evaluation score that would determine their professional assignments.

It was not an extraordinarily difficult system to learn. Many of the teachers said that the Marzano Model listed and coined strategies that they were “already doing” within their own classrooms. However, it was difficult for teachers to reconcile their previous experiences with teacher evaluation with the high-stakes evaluation model that was presented to them in the form of two district-wide meetings. As a result, some teachers were hesitant to support the district’s implementation of this evaluation model simply because they did not understand the district’s expectations of them.

In terms of Habermas’s (1971) domains of human interest, teachers’ responses indicated that the level of emancipatory knowledge was not reached in their experiences with the Marzano Model. Emancipatory knowledge fosters a deep understanding of the information presented. If emancipatory knowledge were reached, teachers would know and understand why the evaluation policies and practices were being implemented and how these policies and practices impact them on the individual and collective levels (Kitchenman, 2008; Prayer, 1993). Teachers perceived a lack of training and support, a lack in clarity of district expectations, and a lack in inter-rater reliability, which ultimately prevented them from attaining a level of emancipatory knowledge with respect to the Marzano Model.

Additionally, Mezirow’s (1981) perspective transformation never took place in its entirety at Silver Creek. Participant responses suggested that teachers were not able to build competence and self-confidence in their newly defined roles as identified by the Marzano Model. Of course, they were still identified as teachers, but the job description changed, in part, due to the evaluation model. The participants revealed that the structure of the lesson plans was
restructured and revised, district-wide professional development shifted from content-driven to process-driven, and a new common language was being put into place to reflect these changes. With time, teachers developed a sense of comfort with the Marzano Model, but their perspectives regarding the Model have not changed quite as dramatically. Adult learning theory suggests that the manner in which the Marzano Model was introduced did not entirely meet the needs of the adult learner.

**Reflective Practice Theory**

Reflective practice involves assessing one’s own perceptions and actions for the purpose of cultivating and growing one’s craftsmanship (Osterman, 1990). The present study reveals that teachers are largely reflective practitioners, seeking ways to grow and develop as highly effective teachers. However, when teachers perceive the Highly Effective rating to be unattainable, it can actually have the opposite effect. Teachers begin to lose interest in continual improvement, externalizing their documented areas for growth as flaws of the “system,” rather than personal characteristics and behaviors that they exhibit in the classroom.

For example, the participants of the present study identified inter-rater reliability as the main source of contention between administrators and teachers with respect to teacher evaluation practices. In spite of the ratings they earned on the proficiency scale, a number of the teachers expressed their beliefs that they were Innovating, regardless of the scores they earned in previous observations. It is imperative that an evaluation tool promotes professional growth through the empowerment it provides teachers. Teachers should feel comfortable with the tool and well versed in its components. The more teachers interact with an evaluation model, the more likely they are to reach a comfort level with the instrument, resulting in increased confidence and, ultimately, reflective practice. If this does not happen, the administration runs the risk of
unintentionally creating an environment in which the evaluation model serves as a scapegoat for all of the performance issues within the organization.

Similarly, it is difficult for teachers to be reflective in their practice if the model they are using as a “growth model” serves a dual purpose. Administrators cannot expect teachers to take risks if their evaluations are at stake. This is a vicious cycle because teachers cannot expect to grow unless they are willing to take instructional risks and attempt new strategies. The majority of the participants were in favor of the Marzano Causal Teacher Evaluation Model as a growth tool, but not as an evaluative tool. Teachers that were part of Project COACH perceived the district’s use of the Marzano Model more favorably, suggesting through their responses that they were largely in favor of an evaluation model that is structured in such a way that administrators are only using the tool to evaluate the teachers at the end of the school year. Teachers that participated in Project COACH reported increased opportunities for professional reflection.

Social Cognitive Theory and Social Cognitive Career Theory

In a qualitative study, the central consideration is perception, how individuals experience or “see” the world in which they live. Social cognitive theory (Bandura, 1977, 1986) indicates that individuals learn from one another. Because humans are social beings, individual perceptions are shaped by their interactions with others. As would be the case in any professional situation, the teachers that participated in this study have constructed their perceptions of their own abilities and professional worth through interactions with other teachers as well as interactions with their administrators. Several participants reported the stress that their colleagues had experienced after the district’s implementation of the Marzano Model in addition to describing their own stress. The social interactions between the teachers help to shape the
opinions of each of the teachers within the district. If teachers develop a negative perception of
the evaluation tool, the tool is less likely to shape their daily practice.

Similarly, the teachers’ perceptions of the proficiency scale and the summative evaluation
ratings influenced teacher efficacy. Teachers expressed concern that they were not Highly
Effective or Innovating, thus fostering a disconnect between administration and teachers as well
as a largely negative initial opinion of the evaluation model. If teachers feel that they are not
capable of being “A” students, they will be less likely to attempt to do so. Conversely, if
teachers feel that they are capable of achieving the desired result, they are more likely to stay
focused and committed to the task at hand. Students can only benefit from teachers that are
engaged as committed and reflective practitioners.

The present study also reveals the value in coaching teachers. Coaching provides
teachers with consistent constructive feedback and instills a sense of efficacy among the
teachers. In general, teachers that participated in Project COACH were more receptive to
feedback and felt more encouraged throughout the evaluation process. According to social
cognitive theory and social cognitive career theory, these positive interactions with
administrators have the power to shape not only teachers’ perceptions of the evaluation model,
but also their perceptions of self-efficacy.

**Recommendations for Practice**

The participant responses are, at times, contradictory and reveal a specific challenge
presented to school leaders when attempting effective evaluation implementation. Wise, et al.
(1984) identified the following characteristics to effective evaluation implementation:
organizational commitment, evaluator competence, teacher-administrator collaboration, and
strategic compatibility. The current study verified the work of these researchers and serves as a
cautionary tale to any school district embarking on a journey of broad-scale organizational change. Change of any kind is a process, not an event, and it requires school leaders to clarify the need for and purpose of the change (Hall & Hord, 2001). In the case of the present study, the change from the “old” evaluation model as a system of three potential nominal outcomes (Satisfactory, Unsatisfactory, Not Using) to the current Marzano Causal Teacher Evaluation Model instilled both clarity and confusion among teachers. The Marzano Model was implemented during the same year that SGOs (Student Growth Objectives) were introduced and teacher tenure reform was taking shape at the state and national levels. Teachers and administrators alike experienced a great deal of organizational changes in the same year that the Marzano Model was introduced. While the district administration has been receptive in soliciting teacher feedback and adapting the evaluation process to suit the needs of the district, these efforts have been overshadowed by the Marzano Model’s first introduction.

This study speaks to the need to recognize that widespread organizational change should be, if at all possible, implemented gradually. Granted, in order to be compliant with AchieveNJ, the school district could not follow a process that reflected the complexity of the change. Successful organizational change takes place in the following three stages: program implementation, evaluation, and operation (Delbecq & Van de Ven, 1971). Change theorists recommend implementing a pilot program before widespread implementation in order to effectively and efficiently evaluate the program prior to a more broad-scale approach. The rapidity with which the district implemented the changeover to the Marzano Model negatively influenced its success in subsequent years because the change took place too quickly. It was difficult for teachers to disassociate their initial experiences with the Marzano Model after the first year in spite of the changes the district has made since. Teachers felt ill prepared for the
change and, as a result, expressed their feelings of inadequacy when initially being evaluated under the model.

As noted multiple times, inter-rater reliability was also a concern due to the subjective nature of teacher evaluation (Herlihy et al. 2014; Peterson, 1987, 2000). Most of the teachers in the current study cited inter-rater reliability as a flaw of the Marzano Model. In order to develop and maintain evaluator competence, it is recommended that districts adopt a specific set of evaluation protocols from which administrators can equitably assess teachers. Dual observations serve to calibrate evaluators to some degree, but engaging in this practice twice a year does not ensure that administrators will increase inter-rater reliability. Districts should be providing annual refreshers on teacher evaluation practices within the district in addition to the communication of the annual adjustments in the state requirements and recommended practices. Along these lines, districts should be mindful of the initial training that is implemented with any broad-scale change. Professional development must be provided initially and consistently thereafter in order to maintain and monitor the status of the change.

It is also recommended that districts consider adopting a coaching approach to evaluation, a practice which Silver Creek has piloted and has experienced success. Teachers expressed their appreciation for more frequent observations throughout the school year with fewer ratings assigned. Understandably, fewer ratings equate to heavier weighting with respect to earned scores; however, teachers that participated in Project COACH were less focused on the numbers and more focused on professional growth, the ultimate goal of teacher evaluation (Hunter, 1980; Marzano, 2012). It is difficult for teachers to engage in an honest, reflective conversation with their evaluators if they know that the perceived shortcomings are being evaluated and rated. Of course, ratings cannot be eliminated entirely, but it is reasonable to
request that districts consider implementing a more comprehensive approach to teacher evaluation.

One final thought regarding the importance of adopting a coaching model: all of the teachers that participated in Project COACH cited the relationship that they cultivated with their evaluators as the single most influential factor in the coaching program. Similarly, teachers that were evaluated under the traditional evaluation model said that they preferred a single evaluator as opposed to multiple evaluators, citing the improved relationship as the primary reason for the preference. Teachers felt that their observations, in addition to their summative evaluations, were more reflective of their professional capabilities when the administrator was able to document a “film” of their year as opposed to a few still frames or snapshots in time through the evaluation process.

The literature indicates that a collaborative and personalized approach to teacher evaluation is most effective (Coleman, 1946; Thompson, 1952). School districts should consider individualizing their observation and evaluation processes. For example, if the teacher is able to zero in on one particular area of growth and wishes to document this growth through multiple measures, she should be afforded the opportunity. Teachers may wish to engage in a video self-reflection of their lessons and submit these to their evaluators. Peer observations, while complex, may be suitable and preferred for some. Teachers may wish to conduct student or parent surveys to assess their effectiveness in a particular area. Teachers may feel that there is more value in maintaining a portfolio to track and assess multiple measures of professional growth. As long as these measures are serving to improve professional practice and, ultimately, student achievement, districts should consider personalizing an otherwise very impersonal process through differentiating their approach to teacher evaluation based on teacher interests.
and readiness levels. If teachers take ownership of the process, they are more likely to own the results.

**Recommendations for Policy**

There seems to be a distinct difference between research-based recommended practices in teacher evaluation and the most widely utilized practices regarding teacher evaluation within schools. The literature supports the idea that evaluation systems that reflect an emphasis on formative evaluation produce higher levels of satisfaction and more reflective practice (Danielson & McGreal, 2000). By and large, however, evaluation policies reflect an emphasis on formal evaluations that tie directly to teachers’ summative evaluation ratings. Walkthroughs are encouraged but not required by any evaluation policy. However, the current study suggests an overwhelming need for administrators to make an effort to engage in more frequent, informal observations in order to both gauge the professional capabilities of teachers and provide the necessary support to foster professional growth.

Policymakers should be mindful of the need for district autonomy with respect to teacher evaluation practices. Teachers should be held accountable for their professional practices and growth but not at the expense of reflective practice. Teacher evaluation is not a “one size fits all” approach. Understandably, minimum requirements across the state must exist in order to maintain inter-district equity. However, equity does not mean providing exactly the same resources in exactly the same context for all school districts. The needs of the teachers within the district should drive the process by which the evaluation model is implemented.

It should, however, be recognized that the state of New Jersey does allow for districts to submit their own evaluation instruments for approval. This factor should be considered when districts consider board policies related to teacher evaluation. If district administrators simply
strive for compliance when make decisions relating to teacher evaluation, they cannot expect much more out of their teachers. The current study reflects a need for differentiation across districts when making decisions regarding evaluation practices. Just as the needs in one district will not mirror the needs in another district, an evaluation tool should not be carbon copied across these districts.

Policymakers and districts alike should also be mindful of the impact that professional ratings have on self-efficacy, job satisfaction, and reflective practice. What does it mean to be Highly Effective? According to state reports, there is a discrepancy amongst districts as to what that truly means. Wide ranges exist between districts regarding the number of teachers rated as Highly Effective in their summative evaluation score. Those in administrative positions cannot possibly think that ranges this wide are mathematically and logically sensible. The current state-approved evaluation models, including the Marzano Model, have helped to bring much-needed specificity to the observation and evaluation processes through the establishment of a detailed framework from which administrators can assess their teachers. Ultimately, an evaluation tool is only as good as the people who are using it. Consistency within as well as across districts would only strengthen the credibility of a teacher evaluation model.

Finally, policymakers should reference change theory before identifying specific timelines for the implementation of broad-scale change. Pilot programs should not be optional, but mandated when implementing such a change. Teachers and administrators would only benefit from increased exposure to an evaluation tool prior to its implementation. As previously noted, change is not an event; education policy should reflect this reality in the form of reasonable timelines and long-range assistance throughout the change process.


**Recommendations for Future Research**

The current study was limited to a single high school in southern New Jersey. Grades K-8 were not considered throughout this study. The teachers that participated in this study frequently mentioned that the Marzano Model might be more suitable for the elementary grade levels. Further exploration into Grades K-8 would provide an enhanced perspective on the perception of the Marzano Model, as these grade levels are assigned an mSGP, a component of their summative evaluation rating that is exclusively derived from student achievement on state standardized test scores. Their perspectives on teacher evaluation could differ slightly considering their evaluations are directly linked to student achievement. In August of 2016, the New Jersey State Department of Education released new adaptations to teacher evaluation requirements. Teachers’ mSGP scores will now count towards 30% of their evaluation score. ESSA also seeks to reward high-performing teachers in low performing schools. In other words, merit-based pay continues to be a discussion in public education. As the stakes become higher in high-stakes teacher evaluation, it is becoming increasingly important to explore teacher perception of evaluation and its impact on teaching practice.

The current study was also limited to the exploration of teachers’ perceptions of the Marzano Causal Teacher Evaluation Model. The perceptions of administrators were not included as part of the research. Qualitative research is primarily focused on the perceptions of lived experiences. Considering the fact that only one perspective of teacher evaluation was explored, future research in exploring administrators’ perceptions of teacher evaluations could add depth to the literature on teacher evaluation.

Additionally, the teachers that participated in this study were all tenured with at least five complete years of experience. This study did not reflect the perceptions of new teachers with
respect to the Marzano Model and evaluation practices. Do new teachers perceive teacher evaluation differently than experienced teachers? Do tenured teachers perceive teacher evaluation differently than non-tenured teachers? The sample of the current study did not allow for me to explore these complex questions.

Finally, the current study focused solely on the Marzano Causal Teacher Evaluation Model. While other literature addressing other state-approved models exists, it is recommended that researchers consider exploring teacher perceptions of other evaluation models in order to further identify best practices in teacher evaluation. It is my hope that further exploration will lead to an increased awareness of the immediate and long-term effects of teacher evaluation implementation and the subsequent changes that directly affect student achievement.
References


*Fixing classroom observations: How common core will change the way we look at teaching.* (2013).


Thompson, E. (1952). So begins – so ends the supervisor’s day. *Educational Leadership, 10*(2), 80-84.


APPENDIX A

2014 Marzano Teacher Evaluation Model Learning Map
2014 Marzano Teacher Evaluation Model

Domain 1: Classroom Strategies and Behaviors
Domain 1 is based on the Art and Science of Teaching Framework and identifies the 41 elements or instructional categories that happen in the classroom. The 41 instructional categories are organized into 9 Design Questions (DQs) and further grouped into 3 Lesson Segments to define the Observation and Feedback Protocol.

Note: DQ refers to Design Question in the Marzano Art and Science of Teaching Framework. The 9 DQs organize the 41 elements in Domain 1.

The final Design Question, DQ10: Developing Effective Lessons Organized into a Cohesive Unit, is contained in Domain 2: Planning and Preparing.

©2014 Robert J. Marzano. Can only be digitized in iObservation. iObservation is a registered trademark of Learning Sciences International®.
APPENDIX B

Letter of Solicitation
Dear Teachers at SCSD:

I am inviting you to take part in a research study. My name is Susan Graziano, and in addition to being an Assistant Principal at Kingsway Regional Middle School, I am also a doctoral candidate at Seton Hall University, College of Education and Human Services. I am conducting a research study in partial fulfillment of the requirements for the degree of Doctor of Education. Upon successful completion of my dissertation, I will receive my doctorate in K-12 Educational Leadership, Management, and Policy.

I have received permission from the superintendent of the school district, Dr. James Lavender, to conduct my research here. The purpose of this study is to examine teachers’ perceptions of the district’s selected evaluation system, the Marzano Causal Teacher Evaluation Model, in particular, the perceived impact the instrument has on teachers’ instructional practices. Examining and exploring teachers’ perceptions of the model is extremely valuable because the information gleaned from this study could provide districts with information that could shape the practices and policies associated with teacher evaluation. The literature suggests that there is a positive correlation between teacher quality and student achievement: As the quality of the teacher increases, student achievement rises. By understanding teachers’ perceptions of teacher evaluation through research, school leaders can identify ways to successfully utilize teacher evaluation models to potentially improve classroom instruction.

I am asking for approximately 15 volunteers, with a range of experience, to participate in this research study. After I have a list of volunteers, I will send out a Demographic Profile Questionnaire to be completed by each volunteer. This questionnaire is very brief and will only require a few minutes to complete. From the teachers’ profiles, I will select 15 teachers that will act as a representative sample population of the Silver Creek High School.

If you are selected from the pool of volunteers to participate in this study, you will be required to sign an informed consent form. I’ll ask you to meet with me for one interview that will be approximately 30 minutes in duration. The interview will consist of questions relating to your experiences with the Marzano Causal Teacher Evaluation Model. You will be asked the same questions as all of the other participants. These interview questions were reviewed by a panel of experts prior to conducting the research. The interview will be recorded with an audio recording device and subsequently transcribed. After your interview is transcribed, you will have an opportunity to review the transcription for accuracy. The interview will be conducted in your classroom or in another area within or outside of the school building, mutually agreeable to both parties.

To participate in this research, it is estimated that each participant must be available for a total of approximately one hour: 30 minutes for the initial interview and 30 minutes for transcription review (if interested). The interviews and review of transcripts will occur over the next few months.

Participation in this study is completely voluntary. You are free to request that your transcript or interview not be used at any time. You are free to withdraw from the study at any point in time. You also have the right to refuse to answer any question that makes you feel uncomfortable. There is no penalty for withdrawing from the study. You will receive no compensation for participating in the research study.
Any information that is obtained throughout this study that can be identified with you will remain confidential and will not be disclosed for any reason. All data collected from you will be coded with a number assigned to you that will not be shared with anyone. Pseudonyms will be used throughout the dissertation. Your real name will never be used. The school district will also be assigned a pseudonym to protect the confidentiality of all personnel at Silver Creek. The identity of all participants will be carefully protected while I am conducting the research and when I report my findings.

All of the research data will be securely stored on a USB flash drive and on the recording device stored in a locked filing cabinet in my home. Once the audiotapes are transcribed and put on USB flash drives, the audiotapes will be erased. I will keep all research data for a period of three years after the research study is concluded. At that point, all data will be destroyed.

I want to sincerely thank you for your consideration to be part of this study. If you are interested in participating in my research study or if you would like to discuss it with me in further detail, please feel free to contact me by phone or email. My phone number is (856) 467-3300 ext. 3010, and my email is graziasu@shu.edu. If you prefer, you may contact my dissertation chair, Dr. Barbara Strobert, at her office at Seton Hall University at (973) 275-2324, or Mary F. Ruzicka, Ph.D., Director of the Institutional Review Board, at (973) 313-6314.

Sincerely regards,

Susan Graziano
APPENDIX C

Demographic Profile Questionnaire
DEMOGRAPHIC PROFILE QUESTIONNAIRE

1. For how many years have you been an education professional?

   _____1-5   _____6-10   _____11-15   _____More than 15

2. For how many years have you been working at Kingsway Regional High School?

   _____1-5   _____6-10   _____11-15   _____More than 15

3. Current age:

   _____20-29   _____30-39   _____40-49   _____50+

4. Sex (Circle One): MALE  FEMALE

5. Highest Degree Attained: BACHELORS  MASTERS  DOCTORATE

6. What grade level/s do you currently teach?

   ____________________________

7. What subject/s do you currently teach?

   ____________________________

8. (Optional): What was your rating during your last summative evaluation?

   _____Highly Effective   _____Effective   _____Partially Effective   _____Ineffective

In order to report statistics regarding the composition of the sample population, I am asking that participants provide specific information regarding their ages and total years of teaching. Please record your exact age and the total number of years you have been teaching below. This information will not be used in any way to identify you as a participant. Thank you.

Age: ___________  Total Number of Years in Education: ___________
APPENDIX D

Informed Consent Form
Informed Consent Form


The Researcher’s Affiliation
The researcher for this study is Susan Graziano. Susan Graziano is a doctoral student at Seton Hall University, College of Education and Human Services.

The Purpose of the Study
The purpose of this study is to explore how teachers perceive the influence of the Marzano Causal Teacher Evaluation Model on their instructional practices. The researcher intends to explore whether or not teachers feel that the Marzano Causal Teacher Evaluation model has guided and/or influenced teachers’ instructional practices. Additionally, the researcher intends to explore teacher’s perceptions of how accurately the Marzano Causal Teacher Evaluation model reflects their professional abilities. Finally, the researcher intends to explore teachers’ perceptions of how, if at all, the Marzano Causal Teacher Evaluation model has influenced teachers’ relationships with their administrators. The participants for this research study will be asked to participate in this study because they are teachers in the State of New Jersey that have had at least one complete year of experience being evaluated under the Marzano Causal Evaluation model. Participation in this research will require approximately 1 (one) hour of the participant’s time to participate in an interview and review the transcription of the interview.

Description of the Procedures
If the participant decides to take part in this study, the researcher will ask each participant to participate in an interview with the researcher for approximately 30 (thirty) minutes. All of the conversation will be audio recorded and then transcribed within two days of the interview. Approximately 15 teachers will be interviewed and asked the same interview questions that were reviewed by a panel of experts prior to conducting the interviews. The interviews will be semi-structured, allowing for the researcher to ask probing follow-up questions, when necessary. The researcher will decide if probing questions are necessary based on the participant’s responses.

The interview will take place after school in the teacher’s classroom with an open door, an available conference room with an open door, or at a mutually agreed upon public place and time. Each participant will be given his/her transcription to read to determine if any modifications or clarifications are needed.

Instruments Utilized During Data Collection
Upon volunteering for the study, participants will complete a Demographic Profile

Department of Education Leadership, Management and Policy
Jubilee Hall • 400 South Orange Avenue • South Orange, NJ 07079 • Tel: 973.761.9377 • Fax: 973.275.2847 • www.sju.edu
Questionnaire, which asks participants to identify information including: number of years teaching, number of years teaching at the research site, current age, highest degree earned, and grade level taught.

During the interview, participants will answer questions that relate to the three guiding research questions.

The following are sample questions that will be asked during the interview:

1. In what ways, if any, do you think your approach to lesson planning and teaching has been influenced by the district’s use of the Marzano Causal Teacher Evaluation Model?
2. What are your thoughts about the accuracy of the Marzano Causal Teacher Evaluation model proficiency scale when evaluating your professional performance?

Voluntary Participation
Participation in this study is strictly voluntary. Any participant may decline to participate, without penalty. You may withdraw from the study at any time, without penalty. If you feel uncomfortable at any point in time, you have the right to decline to answer a question or end the interview at any time, without penalty.

Anonymity
Participants will not be anonymous because the researcher will know the identities of those who volunteer to participate in the study and their responses to the interview questions.

Confidentiality
Participants’ identities will be kept completely confidential. Only the researcher will know the identity of the participants. All participants will be assigned a code number that will be known only to the researcher. The researcher will be the only individual that can identify the participants’ responses to the interview questions. The researcher will not disclose the identity of anyone who participated in the study or the identities of those who contributed individual responses.

Storage of Confidential Data
All documents obtained from the research will be kept in a locked filing cabinet in the researcher’s home. All interviews will be recorded using an audio recording device. All interviews will be transcribed verbatim and saved to a USB flash drive. The USB flash drive, audio recordings, and all printed materials will be stored in the locked filing cabinet when not in use. To protect the anonymity of each of the participants, all audio recordings will be deleted when the participant approves the interview transcript. Each transcript will be identified by the participant’s code number. To protect the participant’s privacy, no names will be used during transcription. The key that identifies the code number for each participant will be stored separately from the data as well. The researcher will keep the data for a period of three years. After the three-year time period
has passed, all data will be destroyed.

Access to Confidential Records
Only the researcher and members of the dissertation committee will have access to data obtained from the participant. The dissertation committee is obligated to protect the data from disclosure outside of the research.

Risk or Discomfort
There are no foreseeable risks involved in participating in this study. You may, however, feel some stress when answering some of the interview questions that pertain to your experiences. You have the right to decline to answer any question that makes you uncomfortable.

Direct Benefits from This Research
There will be no direct benefit to you as a result of participating in the research. However, the information gleaned from this study will add to the literature on teacher evaluation practices. This research could influence school districts when developing, implementing, and sustaining teacher evaluation protocols and policies.

Remuneration
There is no monetary compensation for participation in this study.

Description of Compensation
Since the research study does not include medical treatment, no medical compensation will be offered.

Appropriate Alternative Procedures
The researcher is not aware of any alternative study concerning this issue. The alternative is not to participate.

Contact Information
Please contact Susan Graziano, the principal researcher, with any questions or concerns at (856) 467-3300 ext. 3010. If you have any concerns about this research study that you would like to direct to someone other than the principal researcher, you may contact the Dissertation Advisor, Dr. Barbara Strobert, at her office at Seton Hall University at (973) 275-2324, or Mary F. Ruzicka, Ph.D., Director of Seton Hall University's Institutional Review Board, at (973) 313-6314. If you would like to submit questions or concerns in writing, please use the following mailing address:
College of Education and Human Services
Seton Hall University
400 South Orange Avenue
South Orange, NJ 07079

Audio-tapes
All interviews will be recorded using an audio recording device. Participants will never be identified by name; participants’ code numbers will be used in place of their names. The researcher will personally transcribe all interviews. The tapes and transcriptions will be kept in a locked filing cabinet in the researcher’s home. Once the participant has approved the transcription, the audio-tapes will be erased. Only the researcher will have access to the audio-tapes and the transcriptions. The data collected will be kept for a period of three years after the study has concluded. After the period of three years has expired, all data will be destroyed.

The participant will be given a signed copy of this document for his/her records. One copy will be kept with the research study records.

I have read and fully understood this consent form. I have had the opportunity to ask any questions regarding this consent form. I understand the nature of this study and agree to participate. I also consent to having my interview with the researcher recorded on an audio recording device.

______________________________  __________________________
Signature of Participant             Date
APPENDIX E

Certificate of Completion: Protecting Human Research Participants
Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that Susan Graziano successfully completed the NIH Web-based training course "Protecting Human Research Participants".

Date of completion: 05/06/2016.

Certification Number: 2069556.
Interview Questions

Research Question 1: How, if at all, has the Marzano Teacher Evaluation model influenced and informed teachers’ instructional practices?

1. What does Marzano require and/or look for in terms of instructional practice?
2. How have you used Marzano, if at all, when you develop and implement lessons?
3. What do you like about Marzano, if anything?
4. What don’t you like?
5. In what ways, if any, do you think your approach to lesson planning and teaching has been influenced by the district’s use of Marzano?

Research Question 2: What are teachers’ perceptions of how accurately the Marzano Teacher Evaluation model reflects and captures their professional performance and capabilities?

6. How would you define/describe your experience with each of the proficiency scale ratings (Innovating, Applying, Developing, Beginning) of the Marzano model?
7. What do you like, if anything, about the proficiency scale?
8. What don’t you like?
9. What are your thoughts about the accuracy of the summative evaluation rating(s) (Highly Effective, Effective, Partially Effective, Ineffective) that you’ve earned under the Marzano model?

Research Question 3: What are teachers’ perceptions of the influence of Marzano Teacher Evaluation model on their relationships with administrators?

10. How would you describe the professional conversations between you and your administrator(s)?
11. How, if at all, has the Marzano Model improved the quality of professional conversations between you and your administrator(s)?
12. How would you describe your administrator(s)’ expectations with respect to classroom performance?
13. How, if at all, has the Marzano Model helped to define the expectations of your administrator(s) with respect to classroom performance?

14. What are your thoughts about the feedback you have received throughout the evaluation process under the Marzano Model?

15. How has that feedback influenced your relationship with your administrators?
APPENDIX G

School’s Permission Granted
August 22, 2016

Dear Ms. Susan Graziano:

This letter serves as official notice approving your request to conduct research at Kingsway Regional High School in support of your doctoral requirements at Seton Hall University for your study titled: **Teacher Perception of Marzano Causal Evaluation Model.**

Very truly yours,

Dr. James J. Lavender
Superintendent of Schools

JLL/ss
APPENDIX H

IRB Approval
June 29, 2016

Susan Graziano

Dear Ms. Graziano,

The Seton Hall University Institutional Review Board has reviewed and approved as submitted under expedited review your research proposal entitled “An Exploration of Teacher Perception of the Impact of the Marzano Causal Evaluation Model on Instructional Practices.” The IRB reserves the right to recall the proposal at any time for full review.

Enclosed for your records are the signed Request for Approval form and the stamped original Consent Form. Make copies only of this stamped Consent Form.

The Institutional Review Board approval of your research is valid for a one-year period from the date of this letter. During this time, any changes to the research protocol must be reviewed and approved by the IRB prior to their implementation.

According to federal regulations, continuing review of already approved research is mandated to take place at least 12 months after this initial approval. You will receive communication from the IRB Office for this several months before the anniversary date of your initial approval.

Thank you for you cooperation.

In harmony with federal regulations, none of the investigators or research staff involved in the study took part in the final decision.

Sincerely,

Mary F. Ruzicka, Ph.D.
Professor
Director, Institutional Review Board

cc: Dr. Barbara Strobert
APPENDIX I

Approval for Dissertation Proposal
Candidate, Susan Graziano, has successfully completed all requisite requirements. This candidate’s proposal has been reviewed and the candidate may proceed to collect data according to the approved proposal for dissertation under the direction of the mentor and the candidate’s dissertation committee.

If there are substantive differences between what has been approved and the actual study, the final dissertation should indicate, on separate pages in the Appendix, the approval of the committee for those changes.

Title of Proposed Dissertation:
An Exploration of Teacher Perception of the Impact of the Marzano Causal Evaluation Model on Instructional Practices

Dissertation Committee:

ANTHONY COLELLA, Ph.D.  Signature/Date
Mentor (Print Name)  4/22/16

BARBARA STROBATI, Ed.D. Signature/Date
Committee Member (Print Name)  Barbara Student

Carolyn Satin-Bayly
Committee Member (Print Name)  Signature/Date

__________________________  Committee Member (Print Name)  Signature/Date

__________________________  Committee Member (Print Name)  Signature/Date

Approved by Seton Hall University Institutional Review Board on: ____________________________

Department Chairperson  Signature/Date

Waived by IRB by: ____________________________ on this date ____________________________
APPENDIX J

Lesson Plan Template
Lesson Plan Template

Unit Title:
Standards:

Learning Goal:
Day #:

- Learning Objective(s):
  - Learning Activities: (Include elements from DQ 2, 3, and 4 that you will utilize.):
    - Design Question(s): Elements- DQ2:6, 8, 9

- Learning Assignments:

Assessment(s):

Learning Goal:
- Overall goal of this segment/section of the “unit”
- What do you want your students to understand? What do you want your students to be able to do?
- Tied to standards

Learning Objective:
- More specific aims that lead to the learning goal
- Not activities or assignments- Use taxonomy (Webb’s Depth of Knowledge [DoK])

Learning Activities:
- Guided learning experiences that take place in the classroom; think in terms of Design Questions
  What are you doing to

Learning Assignments:
- Learning experiences designed to be completed *independently* in a class or as homework;
  practice that extends the learning experience

Assessments:
- Measurement(s) used that provides feedback on student progress toward learning goals and how they can improve