

Spring 5-18-2019

An Exploration of Principal Perceptions of the Danielson Framework for Teaching in Providing Feedback and Improving Instructional Leadership

Stephen Hernon
stephen.hernon@student.shu.edu

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

Part of the [Educational Administration and Supervision Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), and the [Educational Leadership Commons](#)

Recommended Citation

Hernon, Stephen, "An Exploration of Principal Perceptions of the Danielson Framework for Teaching in Providing Feedback and Improving Instructional Leadership" (2019). *Seton Hall University Dissertations and Theses (ETDs)*. 2655.
<https://scholarship.shu.edu/dissertations/2655>

AN EXPLORATION OF PRINCIPAL PERCEPTIONS OF THE DANIELSON
FRAMEWORK FOR TEACHING IN PROVIDING FEEDBACK AND IMPROVING
INSTRUCTIONAL LEADERSHIP

STEPHEN HERNON

Dissertation Committee

Barbara V. Strobert, Ed.D., Mentor
Anthony Colella, Ph.D.
Mirella I. Avalos-Louie, Ed.D.

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

Seton Hall University
2019

© Copyright 2019 by Stephen Hernon
All Rights Reserved

SETON HALL UNIVERSITY
COLLEGE OF EDUCATION AND HUMAN SERVICES
OFFICE OF GRADUATE STUDIES

APPROVAL FOR SUCCESSFUL DEFENSE

Stephen Hernon has successfully defended and made the required modifications to the text of the doctoral dissertation for the **Ed.D.** during this **Fall Semester 2018**.

DISSERTATION COMMITTEE
(please sign and date beside your name)

Mentor:
Dr. Barbara Strobert

 12/18/18

Committee Member:
Dr. Anthony Colella



Committee Member:
Dr. Mirella Avalos-Louie



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

Teachers have consistently been proven by research as the most critical factor in the academic success of students (Baker et al., 2010, Darling-Hammond, 2000, Tucker & Stronge, 2005, Taylor & Tyler, 2012). Having a quality teacher therefore provides the best pathway to strong student achievement, and knowing how to evaluate quality teaching becomes critical to student success. A growing body of models for evaluating teachers has formed over the past two decades; and among those are new, research-based models. Principals are key to the success or failure of the new approaches to teacher evaluation because they are in the position to use the evaluation models for formative and summative purposes.

This study sought to understand how principals used *The Danielson Framework for Teaching* to become more effective instructional leaders of their schools in light of the recent implementation of the *Framework* and other responsibilities of a principal.

The results of this study speak to the importance of the proper training of evaluators (principals), the development of school-based cultures of growth and trust, and allocation of financial resources to accomplish on-going training of evaluators and instructional improvement.

Acknowledgments

The completion of this study and my doctoral degree is the work of a multitude of people who have come and gone throughout my life, leaving behind the results of their efforts and wisdom in the form of a better version of myself. Those countless friends, colleagues, and teachers have been a part of my life. Of the many quality teachers I have had since kindergarten, two have been instrumental in contributing to the analytical and research skills I most needed to complete this study: Ms. Joanne Umstot-Verdiglione and Dr. Jeffrey Horn. In high school, Ms. V., and then in college, Dr. Horn, both challenged me with complexity and continually pushed me to unravel that complexity so that I might make sense of it for myself and then justify it for others in a defensible and principled manner. As will be repeatedly discussed in this study, there is no replacement for a highly effective teacher.

Among the many are a few people I would like to thank and acknowledge for their modest, but tireless and impactful, efforts in specifically helping me complete this degree:

- To my mentor, Dr. Barbara Strobert, for your insight, guidance, and motivation. You often referred to me as your "prodigal student." Even though that was meant somewhat in jest, I think it truly demonstrates your commitment as a teacher and mentor in helping your students overcome whatever challenges lie in their path. You truly ensured that I stayed focused and got it done. I can never thank you enough for the extraordinary steps you took to help me finish, and I certainly will never forget.

- To my committee members, Dr. Anthony Colella and Dr. Mirella Avalos-Louie, for helping to push my thinking and my writing to deeper levels. This study is so much more meaningful and clear thanks to the both of you. Dr. Colella, thank you for announcing your retirement, which propelled me to the end!
- To my mother-in-law, Joan Wilton, for the countless offers to watch the children to give me time to work. Thank you for the offers and all the times I accepted those offers.
- To my impetus and catalyst, Dr. Byron Darnell and Mr. Jonathan Morano, for stepping-in to help me when you did not need anything extra to do in your own lives. Byron, countless times in the past several years you have taken time to reach out to me and motivate me to “keep going.” Your texts and phone calls provided more support and motivation than you may think they did – you were the impetus for the completion of this study. Jonathan, you were ready to help at any time; and when I called upon you for that assistance, you never hesitated. Your efforts were tremendous for me – you were the catalyst for the completion of this study.
- To my father, Steve Hernon, because throughout this process you have been there asking, “how it was going,” offering your support and providing unsolicited help. Keeping the right mindset was invaluable for finishing this study, and our talks always helped keep me in the right perspective. For that, and so much more, thank you for all your involvement in helping me complete this study.

- To my mother, Margaret Kelly, who was my first example of a “highly effective teacher.” The fingerprints of your impact on me are on each page of this study. They are invisible, but undisguised. They are between the lines, and in the margins. They are the kitchen table hours of homework assistance. They are the conversations about school. They are the car rides to tutoring and prep classes. Without all that, and more, this would be nothing but blank pages. Thank you for all the love and support.
- To my wife, Jennifer Wilton-Hernon, for everything. There are likely no other wives that would be supportive of their husband beginning a doctorate the same month their twins are born, but that exceptionalism encapsulates everything about you. For these seven years, with unwavering patience and without complaint, you supported me with anything I needed. Thank you for the thousands of hours you let me dedicate to this degree. Thank you for doing everything I noticed that you did, and all those things I did not notice. More than anyone you deserved to lose patience for this to be over, but you never did and that made all the difference. The diploma should have space for two names because without you, mine would not be there. Thank you. I love you.

Dedication

This study is dedicated to Maeve and Liam.

When I was working on this degree, a common saying among the professors was, “The dissertation is not your life’s work.” I have been fortunate to attend great schools, with great teachers, and with great friends. My wish is that you, too, have great learning and social experiences on your own paths. My life’s work is to help you have that.

Love always.

Table of Contents

Chapter I. INTRODUCTION	1
Context of the Problem	1
Statement of the Problem.....	5
Purpose of the Study	8
Research Questions.....	8
Theoretical Framework	9
Design and Methodology.....	9
Significance of the Study	11
Limitations.....	11
Delimitations.....	13
Definition of Terms.....	13
Summary	14
Chapter II. REVIEW OF THE LITERATURE	16
Introduction.....	16
Literature Search Procedures	16
Criteria for Inclusion and Exclusion of Literature	17
An Integrated Historical Evolution of Teacher Evaluation & The Field of Evaluation.....	18
Pre-Tylerian Period: Before 1930	19
The Tylerian Period and the Age of Innocence: 1930-1957.....	21
The Age of Realism: 1958-1972	23
The Age of Professionalism: 1973-2007	27
The Age of Accountability: Teacher Evaluation in the 21 st Century	34
<i>The Framework for Teaching</i>	37
Theoretical Frameworks	41
Michael Scriven	41
Fullan Leading Learning.....	43
Summary.....	46
Chapter III. METHODS	47
Introduction	47
Background.....	47
Design	52
Profile of the Site.....	53
Sampling	56
Data Collection	57
Data Analysis	62
Protection of Human Subjects and Ethical Consideration.....	64
Validity and Reliability	65

Role of the Researcher and Researcher Bias.....	67
Summary	69
Chapter IV. RESULTS	70
Introduction.....	70
Themes Related to Research Question 1.....	71
Introduction to the Findings Related to Research Question 1.....	71
Differences from Past Experiences	72
Self-directed Preparation and Training	73
Variability in Ratings and Relevance	75
Summary of Findings for Research Question 1	78
Themes Related to Research Question 2.....	78
Introduction to the Findings Related to Research Question 2	78
Useful Tool in Providing Formative Feedback and Coaching	79
Focused Conversations about Teacher Practices	81
Relationships and Personalities	82
Summary of Findings for Research Question 2.....	84
Themes Related to Research Question 3	84
Introduction to the Findings Related to Research Question 3	85
Clear Expectations for Job Performance	85
Evidence to Determine Teacher Tenure	86
Summary of Findings for Research Question 3	89
Themes Related to Research Question 4	89
Introduction to the Findings Related to Research Question 4	89
Improving Instructional Leadership Practices	90
Setting High Expectations.....	93
Principals as Leading Learning	94
Summary of Findings for Research Question 4.....	96
Summary	97
Chapter V. CONCLUSIONS AND RECOMMENDATIONS	98
Introduction	98
Research Questions.....	100
Summary and Discussion of Major Findings.....	100
Findings Related to Research Question 1.....	100
Findings Related to Research Question 2.....	102
Findings Related to Research Question 3.....	104
Findings Related to Research Question 4	106
Findings Related to Theoretical Frameworks	107
Scriven’s Formative and Summative Evaluation	107
Fullan’s Leading Learning	108
Recommendations for Policy	110
Recommendations for Practice.....	111
Recommendations for Future Research.....	113

REFERENCES	116
APPENDICES	129
APPENDIX A: District’s Permission Granted	129
APPENDIX B: IRB Approval	130
APPENDIX C: Certificate of Completion: Protecting Human Research Participants.....	131

List of Tables

Table 1. Seven Fountains' 2016 New York State Exam Percent Proficient	54
Table 2. Seven Fountains' Observation Number Requirements	55
Table 3. Summary of Demographic Profile Questionnaire for Each Participant.....	57
Table 4. Research Questions and Corresponding Interview Questions	58
Table 5. Sample Interview Questions Derived from Theoretical Framework	60
Table 6. Preliminary Set of Data-Driven Codes	63
Table 7. Emerged Themes as Related to Research Questions	64

CHAPTER I

INTRODUCTION

The school principal is at the forefront of drastic changes impacting the teaching profession. The American educational system has long been fueled by an urgency for improvement. From *A Nation at Risk* (1983), to *No Child Left Behind* (2001), to *Race to the Top* (2009), the urgency grew, bringing tectonic shifts in policy (specifically in curricula standards, school choice, and teacher evaluation) that impacted practice. International data from *Trends in International Mathematics and Science Study* (TIMSS) and *Programme for International Student Assessment* (PISA) accelerated calls for the urgent need for change and results. There were many calls for American schools to increase overall student achievement and focus especially on closing the achievement gap between minority and majority students. For policymakers, the stakes are high because American businesses need qualified employees to continue America's financial dominance in the face of rising international competition from Western and non-Western businesses alike. For educators, the stakes are high, as calls for greater principal and teacher accountability have brought changes to how teaching is measured and how teachers are employed. Positioned at the center of actions to improve teaching and teacher quality, principals use new tools to help make teaching excellent in every classroom.

Context of the Problem

The reality of the teacher evaluation culture in 2017 is one shaped by the convergence of four forces: recent research on quality of teaching, philanthropic involvement in education, policymakers' efforts to change laws governing teacher

evaluation, and pressure to remove poor-performing teachers (Editorial Projects in Education Research Center, 2015).

The teacher has consistently been proven by research as the most critical factor in the academic success of students (Baker et al., 2010, Darling-Hammond, 2000, Tucker & Stronge, 2005, Taylor & Tyler, 2012). Having a quality teacher therefore provides the best pathway to strong student achievement, and knowing how to evaluate quality teaching becomes critical to student success. A growing body of models for evaluating teachers has formed over the past two decades; and among these are new, research-based models created by researchers including Charlotte Danielson, Kim Marshall, and Robert Marzano. Despite more, and deeper, ways to evaluate teachers, early and recent indications after their implementation yielded teacher ratings similar to the levels that existed before implementation of the new models (Keesler & Howe, 2012; Barge, 2012; Sawchuk, 2013; National Council on Teacher Quality, 2017). A few studies have provided evidence that the new evaluation models could provide improvement in teacher quality (Dee & Wyckoff, 2013; The MET Project, 2012). Despite debate on the impact of the models, the well-established importance of an effective teacher on student achievement fuels an iterative rollout of the new evaluation models. Being in the position to use the evaluation models for formative and summative purposes, principals are key to the success or failure of the new approaches to teacher evaluation.

Bill Gates is likely the most influential non-educator on education in a generation. Philanthropy and education policy have a long partnership, but that partnership is changing because, as Reckhow states, the 21st century philanthropic foundations, such as

the Bill & Melinda Gates Foundation, are capitalizing on the convergence of the following:

1. the growing role of the federal government in education
2. the expansion of market-based reforms
3. the changing nature of philanthropy” (Reckhow, 2013)

The 21st century’s philanthropists are venture philanthropists who aim to use their investments in education to yield targeted results and be able to expand to many schools across the country and the globe (Reckhow, 2013; Saltman, 2010; Scott, 2006). People like Bill Gates, Mark Zuckerberg, and Warren Buffet are releasing millions of dollars into education policy. Zuckerberg’s \$100 million investment in Newark’s public schools, in his mind, aimed to be a “blueprint for national replication across America’s urban centers to transform the lives of its youth” (Russakoff, 2015). Venture philanthropy focuses largely on school choice (mostly the creation and expansion of charter schools) and teacher evaluation/quality/effectiveness. For example, in the Measures of Effective Teaching Project (or MET Project) the Gates foundation invested \$45 million to investigate, with over a dozen other organizations, “how evaluation methods could best be used to tell teachers more about the skills that make them most effective and to help districts identify and develop great teaching” (The Bill & Melinda Gates Foundation, 2010). In the past 15 years, the accelerated flow of money to advocacy and action coupled with the attention brought to teacher evaluation by prominent, famous American venture philanthropists has driven teacher evaluation into the spotlight.

On July 24, 2009, President Barack Obama said, “If you put outstanding teachers at the front of the classroom,” a state could qualify for some of the \$4.35 billion grant

funding made available through the Race to the Top (The White House, n.d.). Race to the Top (RTTT) was the largest federal investment in public education ever. The specific criteria for the award expected that applicant states would propose a plan that included “revising teacher evaluation, compensation, and retention policies to encourage and reward effectiveness” (The White House, n.d.). The impact of RTTT has been significant on teacher evaluation. In 2009, just 15 states required annual teacher evaluations, but by 2016 forty states had implemented teacher evaluation systems (and those even included student achievement data as a portion of the evaluation) (National Council on Teacher Quality, 2011, 2016). In just eight years, the federal government and states have initiated massive overhauls of teacher evaluation systems.

The Race to the Top grant scoring criteria reveal no larger aim than improving the quality of teaching. The largest component of the scoring, at 28%, was determined by “Great Teachers and Leaders, and the largest subgroup of those points was targeted at changing the use and strength of evaluations to, among other things, prevent tenure and remove tenure from teachers (U.S. Department of Education, 2010). Teachers represent the most significant organized labor force in the United States; and during the Great Recession, collective bargaining and its benefactors came under intense scrutiny from budget-strapped governors and the populace. A 2009 Gallup poll listed the support for unions at less than 50%, which was the lowest ever recorded (Gallup, 2009). Teachers being the largest portion of unionized workers became labelled as what Governor Tim Pawlenty of Minnesota called “exploiters” (Surowiecki, 2011) and emboldened Governor Chris Christie of New Jersey to state, “The time to eliminate teacher tenure is now” (D’Amico, 2011). As reported by The New Teacher Project in *The Widget Effect*, nearly

all teachers in their study of 12 districts were rated satisfactory, which is consistent with other findings (Weisberg, Sexton, Mulhern, & Keeling, 2009). These satisfactory ratings fuel public and policymaker dissatisfaction with teachers when viewed in the context of student results on international tests such as PISA and TIMS. According to some, the United States has not fared well on the PISA or TIMS exams, in comparison with other developed countries, and that continued through the 2015 administration of the exam (Organization for Economic Cooperation and Development, 2016; Strauss, 2016). Others, such as Tienken (2013), rebut the findings that American students are not performing adequately on international benchmarks. Despite Tienken's rebuttal to the growing narrative about America's failing schools, considering the combined financial stress of the Great Recession and the rising anti-labor sentiment with the decades long narrative of the failing American school system, the conditions for reformers to push for large-scale changes in teacher evaluations were excellent.

Beginning in earnest in the mid-to-late 1980s, the push for curriculum standards and school choice options provided potential instructional and organizational solutions to the urgency; but by the Great Recession of 2008, the spotlight swung decisively to focus on improving teachers. Policymakers, venture philanthropists, and the public pushed for change, and the research-based evaluation models provided the tools to help ensure that every child had a great teacher. States, districts, and building supervisors were charged with using these tools to improve or remove each and every teacher in America.

Statement of the Problem

Educational leaders, at all levels, are faced with challenges to improve every teacher's ability to help all students learn and achieve more. They need effective tools to

provide formative and summative feedback for teachers. These tools need to be able to both work with teachers to improve their impact on student learning and provide performance ratings to help determine overall effectiveness. Ensuring that all classrooms have excellent teachers could yield tremendous growth for American students because research proves that the teacher is the most significant in-school factor in a student's learning (Baker et al., 2010; Darling-Hammond, 2000; Tucker & Stronge, 2005; Taylor & Tyler, 2012). As important as teachers are to student learning, principals are the second most significant factor (Leithwood, Louis, Anderson, & Wahlstrom, 2004).

For many decades, the system for ensuring teacher excellence varied considerably from school to school, district to district, and state to state. In many instances teacher effectiveness was measured by teacher behaviors that likely had no impact on student learning. By the 1960s, teacher classroom observations became the core of teacher evaluation. Principal observations of teachers are the method used most frequently (Brandt et al., 2007). The general components of the principal evaluation of teachers included a pre-observation conference, the observation, post-observation conference, and a written report (Heneman, Milanowski, Kimball, & Odden, 2006). During the 1990s push for student learning standards, new concepts in teacher evaluation began to take form that also determined a standard of expectation. In 1996, Charlotte Danielson published *The Framework for Teaching* as a researched-based tool for teachers to better hone their teaching craft in the domains of Planning and Preparation, Classroom Environment, Instruction, and Professional Responsibilities (Danielson, 2013). The Danielson *Framework* is among the most used evaluation tools in today's school districts

that helps to create a common language of, and standards of expectation for, effective teaching regardless of the state, district, or school.

Considering the research proving the importance of teachers and principals on student achievement, these new teacher evaluation models, such as the Danielson *Framework*, are potentially the most important new development in education in decades. Educational technology, charter schools, Common Core State Standards, and other new developments have not been proven as pivotal to student learning as teacher effectiveness. Therefore, the specific tool and how that tool is used by principals to improve instruction and determine continued employment is essential and consequential. With models like the Danielson *Framework*, principals are now challenged to utilize teacher evaluation tools for both formative and summative purposes. It is no longer enough to use an evaluation instrument simply for decisions of tenure and termination. The evaluation instrument must help principals work with teachers to improve their instruction throughout their career so that all students can achieve excellence. Teacher effectiveness, as measured by observations, has been shown to improve student achievement (Gallagher, 2004; Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004), but principals must be using an effective tool in an effective way in order to develop and improve each teacher's abilities throughout the teacher's career. Principals need specific and actionable data to provide the individualized supports that a teacher needs to continually improve practice and the potential impact on student learning. Policymakers and the public have great faith that these new teacher evaluation models will provide principals with the tools needed to positively affect teachers' skills (U.S. Department of Education, 2009).

Purpose of the Study

The purpose of this study was to examine principals' perceptions of their ability to effectively use Danielson's *The Framework for Teaching* to provide formative feedback to teachers, make summative decisions on tenure and termination, and improve their instructional leadership of the school.

Changes and added complexities to teacher evaluation represent a significant shift in the principal's responsibilities. Although this shift began in the late 1980s, the recent decade has seen a dramatic acceleration in combination with a considerable array of non-instructional, but still vital, principal responsibilities (Fullan M. , 2014). This study sought to understand how the principals used *The Framework* to become more effective instructional leaders of their schools in light of the recent implementation of Danielson's *Framework* and other responsibilities of a principal.

Research Questions

1. How do principals perceive their preparation to utilize the Danielson *Framework*?
2. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is effective in providing formative feedback to teachers?
3. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is beneficial to the tenure and termination processes?
4. To what extent, if any, do principals perceive that the use of the Danielson *Framework* has improved their instructional leadership?

Theoretical Framework

The Danielson *Framework* provided a conceptual and specific view of quality teaching to shape the study because it serves as “the foundation for professional conversations among practitioners as they seek to enhance their skill in the complex task of teaching” (The Danielson Group, 2013). The guiding theories for this research were Michael Scriven’s theories on evaluation (more specifically formative and summative evaluations) and Michael Fullan’s research on the principal as “lead learner” (Fullan, 2014). The Scriven evaluation theories framed my view of both the organizational level and individual levels of the practices of evaluation, including the position in which the principal was both the formative and summative evaluator. Fullan’s lead learner provided the researcher with a, perhaps idealistic, lens to help understand the introduction of the *Framework* in light of effective principal instructional leadership.

Design and Methodology

After receiving permission from the superintendent of the Seven Fountains School District to conduct the research, I interviewed 12 principals, all with more than two years as a principal within the system. The participants were chosen purposefully. The school system was chosen because of its size, location, and recent transition to the Danielson *Framework*. The school system comprises 25 PreK-8th Grade schools with over 6,000 students that transitioned to the use of the Danielson *Framework* for teacher evaluation beginning in November of 2015. The school system is of average academic performance (as measured by New York State Common Core Exams), includes both urban and suburban school communities, and contains sizable teacher and principal populations. My research began 24 months after the transition to the Danielson *Framework* evaluation

model. This timing, in this school system, provided me the data necessary to address the research questions comprehensively. In addition, the timing within the national context was important given the heightened focus placed on teacher evaluations in the past decade.

This study was best conducted using qualitative methods. In seeking to understand the experience of particular principals newly working with the Danielson *Framework*, qualitative methods would help uncover and bring to life this unique phenomenon (Merriam, 2009). While the Danielson *Framework* does deconstruct teaching quality into specific indicators, the qualitative method helps in analyzing the use of the *Framework* because, as stated by Coughlin and Cronin, “Qualitative research asserts that a phenomenon is more than the sum of its parts, and must therefore be studied in a holistic manner” (Coughlan & Cronin, 2007).

Semi-structured interviews were conducted and recorded as part of the data collection process. Interview questions were designed from a review of the literature and in consultation with a group of experts. Questions were intentionally designed and delivered in the same format to insure the credibility and authenticity of the interview process. Interviews were recorded, transcribed, and coded in an attempt to understand themes and/or patterns that emerged from the data. The interview questions were designed to dig deeper into the principals’ experiences with the Danielson *Framework*. In order to establish a respectful and truthful rapport with the interviewees, the research guaranteed the privacy of the participants. Pseudonyms were used for all participants and any identifications of schools in the school system. Every effort was made to ensure the

data from the interviews were safe. The data records included interview notes, printed papers, and cassette recordings of the interviews.

Significance of the Study

Teachers and principals represent the most significant, controllable variable on student achievement (Baker et al., 2010; Darling-Hammond, 2000; Tucker & Stronge, 2005; Taylor & Tyler, 2012). This study therefore focused on the critical components of student achievement and how Danielson's *Framework* contributes to those components through the evaluation process from the perspective of principals. Principals' ability to measure and help improve teaching is critical. As Stufflebeam and Shinkfield (2007) stated, "Society has a critical need not only for competent evaluators but for evaluation-oriented decision makers as well" (p. 5).

The ability to continually support teachers throughout their careers and to identify teachers who could serve as coaches or other instructional leaders to their peers would be invaluable to schools and to the field of education (Darling-Hammond & Ducommun, 2010). Ample research supports the use of classroom observations to support instructional improvement (Gallagher, 2004; Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004). Since classroom observations comprise a significant portion of *The Framework*, it is important to publish findings that may explain why early data from the new evaluation models, including *The Framework*, have not seen much change in teacher ratings (Keesler & Howe, 2012; Sawchuk, 2013; National Council on Teacher Quality, 2017). The research literature on the principals' perspectives on the use of the Danielson *Framework* to improve instruction is limited, and it would be beneficial to better understand this seemingly disjointed phenomenon.

Limitations

The limitations of this study are as follows:

1. I interviewed principals from one urban and suburban school district in southern New York State. The findings of this study are limited to the perspectives of these particular participants and therefore may not represent the perspectives of a large population of principals.
2. School districts within New York State can choose among a wide variety of teacher evaluation models, and this one school district was utilizing Danielson's *Framework for Effective Teaching*. This study is therefore only representative of the Danielson model and not of the many other models that may improve instruction and instructional leadership.
3. The sample size of 12 principals represents 48% of the school principals within the researched district. While not even representing half of the total population, this is a limitation of the study.
4. The use of interviewing for data collection is a limitation because the researcher must assume that the participants are truthful in all responses.
5. In a qualitative study the researcher is the instrument of data collection and analysis. My own knowledge, experience, and perspective were therefore present during the research and they are likely to have influenced the results of this qualitative study.
6. While I do not work, and have never worked, within the district of study, I know many principals within the district professionally and worked in a related district

under the same leadership and governance. This could have influenced participants' responses during interviews.

7. It is possible that the study attracted principal participants with strong feelings, either way, about the Danielson *Framework* or teacher evaluation in general. This could have impacted the findings of the study.

Delimitations

The delimitations of this study are as follows:

1. This study included principals with at least two years of experience within the district.
2. This study was limited to one, non-public school district within southern New York State.
3. This study was limited to principal perspectives of the Danielson *Framework for Effective Teaching*.
4. This study was limited to the perspectives of principals without researching the impacts of the *Framework* on the teachers or students.
5. This study had a sample of 12 participants, and a small sample may not be applicable to larger districts.
6. This study was limited to a suburban district and may not be applicable to urban districts.

Definition of Terms

The following terms were used throughout this study:

Ineffective. A term used in Danielson to indicate a rating. Ineffective is the lowest of four ratings and indicates the teacher is not adequate in a specific area of

teacher practice.

Developing. A term used in Danielson to indicate a rating. Developing is the second lowest of four ratings and indicates the teacher is somewhat adequate in a specific area of teacher practice.

Effective. A term used in Danielson to indicate a rating. Effective is the second highest of four ratings and indicates the teacher is proficient in a specific area of teacher practice.

Formative Feedback. Feedback provided during the formation of a product or service with the goal of improving the final product or service.

Highly Effective. A term used in Danielson to indicate a rating. Highly Effective is the highest of four ratings and indicates the teacher is beyond proficient in a specific area of teacher practice.

Instructional Practices. Observable teacher actions based on a known set of research- or theory-based best practices.

Rating. A score provided to teachers based on the Danielson observation rubric. There are four possible ratings: Ineffective, Developing, Effective, and Highly Effective (lowest to highest).

Summative Feedback. Feedback provided at the conclusion of a product or service with the goal of comparing quality.

Summary

This study illuminates the experiences of 12 principals transitioning to one of the new teacher evaluation models, the Danielson *Framework for Teaching*. In light of the shifting realities of educational leaders over the past two decades, the findings of this

study may offer valuable insight for instructional leaders at all levels. Evaluations should help identify and foster good teaching. Good teaching on a large scale can transform individuals lives; and as Michel Scriven stated, principals should consider “the extent to which evaluation has made a contribution to the welfare of humankind” (Stufflebeam & Shinkfield, 2007, p. 370).

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

This chapter provides a review of the existing literature pertaining to teacher evaluation and principals' perception of the effectiveness of the evaluation process. The chapter begins with a discussion of the nature of evaluations. This section is followed by an integrated historical evolution of teacher evaluation and the field of evaluation, which examines the impact that the historical progress of education and the development of the field of evaluation had upon each other. The literature review continues with a section on Michael Fullan's leading learner theoretical perspective. This is followed by a discussion of the theoretical perspectives of Michael Scriven's consumer-oriented evaluation. The literature review concludes with an analysis of Charlotte Danielson's *Framework for Teaching*. The examination of this literature served to frame the problem of this study and guide the formation of the research questions.

Literature Search Procedures

The review of the literature related to understanding principals' perspectives on the effectiveness of the Danielson *Framework* to improve instruction was conducted using a number of resources found in the Seton Hall University library, Academic Search Complete, Dissertation Abstracts International (DAI), EBSCOhost Research Databases, Educational Resource Information Center (ERIC), ProQuest Social Science Journals, ProQuest Multiple Databases, Google Scholar, New York Public Library's General Research Division at the Stephen A. Schwarzman Building, United States Department of Education, and ProQuest.

Search terms that were entered into multiple databases, and in multiple combinations, included teacher evaluation, Danielson *Framework*, teacher effectiveness, evaluation, feedback from evaluations, professional learning, Race to the Top, No Child Left Behind, A Nation at Risk, standards-based evaluation systems, history of teacher evaluation, teacher evaluation policy, lead learner, Michael Fullan, Michael Scriven, and Charlotte Danielson.

Most of the works reviewed and included were peer-reviewed scholarly publications within the past 15 years; however, some works dated outside this time frame were included to provide the proper historical perspective and contribute to my understanding of evaluation and principals' perception of teacher evaluation. The style guidelines used in formatting this dissertation were obtained from the *American Psychological Association Manual*, 6th edition (2010).

Criteria for Inclusion and Exclusion of Literature

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

- English language 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 education research journals
- Books and book chapters on teacher evaluation and evaluation theory and practice

- 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

Studies were considered for exclusion from the review if the following criteria were met:

- Literature relating to the perceptions of higher education professionals regarding evaluation of their, or their subordinates, effectiveness;
- Literature that is not written in English; and
- Literature that was conducted in non-public schools in the United States.

An Integrated Historical Evolution of Teacher Evaluation & the Field of Evaluation

The understanding of good teaching has been an ever-changing concept influenced most strongly by evolving educational research and political pressures (Clemetson, 2000). The culmination of these influences formed the purposes for teacher evaluation today, which according to Haefele (1993) are as follows:

- Screen out unqualified persons from certification and selection processes
- Provide constructive feedback to individual educators
Recognize and help reinforce outstanding services
- Provide direction for staff development practices
- Provide evidence that will withstand professional and judicial scrutiny
- Aid institutions in terminating incompetent or unproductive

personnel

- Unify teachers and administrators in their collective efforts

to educate students (1993, p. 7).

Pre-Tylerian Period: Before 1930

Ralph W. Taylor is often spoken of as the father of educational evaluation; but prior to his writings and influence, evaluation in education did exist. From the mid 19th century to the turn of the 20th century, educators, researchers, and city governments began experimenting with how to evaluate student abilities and school effectiveness. By the turn of the 20th century the concept of education itself was dominated by two conflicting views: those of Edward Thorndike and John Dewey.

Educational evaluation in the 19th century was dominated by surveys of schools. Growing from the early 19th century practice of oral examinations, by mid-century, schools in the Northeast, specifically in Boston, began written examinations of students. Horace Mann championed this use of written exams in Boston to serve as evaluative measures of a school's effectiveness in helping students to learn; and by the end of the century, end-of-term exams became commonplace. The Boston survey (1845), Joseph Rice's survey in New York City (1885), and the Cleveland Education Survey (1915) were among the very first large-scale evaluations of anything education-related in the United States and were considered by many to be strong and reliable ways of determining the quality of education. As with current debates regarding the use of student achievement data in the evaluations of schools (and thereby teachers in those schools), the use of surveys and end-of-term style examinations were contested by educators (Stufflebeam & Shinkfield, 2007).

Scientifically analyzing survey data in Boston, New York City, and Cleveland helped open the door to other avenues of school evaluation, especially if those were scientifically and data-based, of school evaluation. In the early 20th century, school leaders began adapting Frederick W. Taylor's scientific management principles to schools. Taylor had successfully developed and applied his principles of efficiency and standardization to business (mostly factories and mines) (Taylor, 1911). Galvanized by the successes of scientific management in the business world, a convergence of education leaders, politicians, and researchers backed expansion of scientific management into education. Potentially the father of educational standardization, Edward Thorndike spearheaded the implementation of standardized tests to students across the country stating that the evidence would help determine efficiencies and show quality in education (Stufflebeam & Shinkfield, 2007).

In the early 20th century, changes in education extended beyond student standardized testing directly to teacher evaluation. Thorndike, along with Cubberley and Wetzel, embodied the new industrial view of education. The concept of schools as factories stemmed from Frederick Taylor's principles of scientific management, whereby the approaches to effectiveness of Taylor's business management were applied to teacher and school evaluation. Evaluation consisted of observable features and actions in the classroom, and the use of student data to inform the teacher about student abilities and skills. As with scientific management, having the correct series of actions for the teachers would yield the most learning for the students. Principals, and other supervisors, filling the business world role of "managers," were expected to ensure, through observation of practice, that teachers followed the correct, essentially prescribed, actions known to yield

student learning and skill acquisition (Wetzel, 1929; Cubberley, 1916; Taylor, 1911). Describing the ideal educational environment, Cubberley stated that schools are “factories in which raw products (children) are to be shaped and fashioned into products to meet the various demands of life” (p. 338). Proponents of the industrial view of education saw the role of teachers to do the “shaping” and “fashioning”; or, to continue the metaphor, teachers were to act as “machines,” efficiently, consistently, and uniformly instructing students. The prominence of Cubberley, Thorndike, and Wetzel’s writings throughout the majority of the first half of the 20th century and into the second half, can still be seen influencing the teacher evaluation systems of today (Ellet & Teddlie, 2003).

The Tylerian Period and The Age of Innocence: 1930-1957

Ralph W. Tyler is the father of educational evaluation, not simply because he coined the term “educational evaluation” but because more than any other individual, he influenced the concept of evaluation within the context of schools from the early 1930s well into the 1960s. His ideas are still felt but have been added to greatly. Tyler’s conceptualization and definition of educational evaluation are simple: the extent to which objectives are achieved. The behavioral objectives of Tyler and behavioral psychology tools employed by Dewey coalesced as *progressive education*. Tyler’s and the progressive education movement’s influence was due largely to his ability to study and publish his theories through his work as lead researcher for The Eight Year Study (Stufflebeam & Shinkfield, 2007; Tyler, 1949).

Tyler’s research on objectives and outcomes had wide-ranging impacts both in education and non-educational fields. His approach did not require comparison/experimental groups because the approach was not intended to compare to other subjects.

Additionally, in contrast with the survey approach of the earlier decades before Tyler's writings, this "approach concentrates on direct measures of achievement, as opposed to indirect" measures (Stufflebeam & Shinkfield, 2007, p. 36) such as quantifying program inputs (Stufflebeam & Shinkfield, 2007; Tyler, 1949).

In the first few decades of Tyler's influence on educational evaluation, some changes in teacher evaluation toward the objective/outcome approach occurred, but also much of teacher performance was determined by the teacher's character. Teacher personality traits, known as "presage variables," were determined by educators in the 1940s and 1950s to be important for teacher effectiveness. Despite the use of these variables in the evaluation of teachers' effectiveness, there is virtually no research evidence to support that presage variables help student learning and teacher effectiveness (Ellet & Teddlie, 2003; Danielson & McGreal, 2000; Kratz, 1896).

Prior to the 1950s, teacher evaluation had three major approaches that served as competing theories of teacher evaluation: an "industrial" view, an "objective/outcome" view, and a "teacher traits" view. The battle of ideas between the industrialists Thorndike, Wetzel, and Cubberley and the progressivists Tyler and Dewey brought many changes to the educational evaluation of students and teachers by the end of the 1940s. To varying extents these three viewpoints can be seen in present day teacher evaluations, but the industrial view clearly dominates the field and practice of teacher evaluation in the present (Marzano, Pickering, & Pollock, 2011).

The years that encompass 1930-1957 reflect a dichotomy in educational evaluation and improvement. From 1930 until the end of the Second World War, educational researchers provided new, rich research from a variety of perspectives that

contributed to the growing body of educational evaluation literature and practices. After the Second World War ended, much of that growth slowed to nearly imperceptible levels. As Stufflebeam and Shinkfield explained, educational systems grew considerably in the decade immediately after the Second World War; but during this time, which they named The Age of Innocence, “society had no particular interest in holding educators accountable, identifying and addressing the needs of the underprivileged, or identifying or solving problems in the educational system” (2007, p. 36).

The Age of Realism: 1958-1972

Prior to 1958, most school decisions were made entirely at the local level, but that would change as significant foreign and domestic issues would cause the federal government to become more deeply involved in the education of the country’s children. On September 2, 1958, when President Dwight Eisenhower signed the National Defense Education Act into law. The law catapulted the federal government into State education funding, and in effect policy and practice as a response to heightened concerns over the Soviet Union’s perceived strengths in math and science that were demonstrated with the launch of *Sputnik*. Through this action, the federal government declared an emergency need for improvement in education:

The push to enhance basic skills acquisition and improve science and mathematics teaching encouraged research into what teachers did or could do to improve basic skills. This time period coincided with significant advances in supervision skills and classroom observation techniques. (Danielson & McGreal, 2000, p. 13)

With this surge of the federal government's and national public's interest in education came a strong interest in evaluation of student achievement, program effectiveness, and teacher effectiveness. In the years immediately following the National Defense Education Act, four evaluation methods emerged as prominent methods for assessing quality and value in education: Tyler's objective versus outcome method, standardized testing of students, professional judgment, and field experiments. It was in light of all these studies and evaluations that were not generating benefit for school districts and teachers that Cronbach challenged the methods of evaluation being used. Cronbach, who studied with Tyler, argued that the evaluators focus more on providing information to practitioners that could help to guide (Cronbach, 1963).

While concerns for the quality of the American versus the Soviet education system focused the federal government and general public on curriculum and overall achievement, domestically socioeconomic inequities (specifically income and race) in achievement led to the creation of the Title I program of the Elementary and Secondary Education Act of 1965. Since the aim of Title I was to help disadvantaged students, the statute was specific in requiring school districts to evaluate the effectiveness of the program and teaching. Districts used a diluted Tylerian approach by assessing objectives and outcomes, not doing so in a local way but in a large-scale manner that Tyler would argue to be contrary to his findings (Stufflebeam & Shinkfield, 2007).

The intervention of the federal government in local education led to an abundance of funding and political interest. The plethora of evaluation methods used yielded little practical assistance. Harris argued that the work of principals and other educational supervisors was as follows:

Characterized by very diverse human relationships, a multiplicity of kinds of tasks, and no fixed locus of operation. The supervisor works in many organizational climates, deals extensively with subordinates, peers, and superordinates, ranges over a wide variety of substantive and procedural problems, produces no readily visible product, is held only vaguely accountable for certain ongoing events in the school, and is almost immune to systematic evaluation.

(Harris, 1965, p. 129)

By the end of the 1960s and early 1970s, many new methods and theories of evaluation were emerging. Some by Provus (1969), Hammond (1967), Eisner (1975), and Metfessel & Michael (1967) built on the Tylerian method of objectives versus outcomes evaluation. Others such as Glaser (1963), Tyler (1967), and Popham (1971) focused on standardized assessment and the movement away from norm-referenced to criterion-referenced testing. Some researchers, including Scriven (1967, 1974), Stufflebeam (1967, 1971), and Stake (1967), pioneered new methods and a serious reconceptualization of evaluation.

Robert Stake built upon Tyler's objective/outcome evaluation model, "calling for examination of background, process, standards, and judgments as well as outcomes" (Stake, 1967, p. 85). For example, in Stake's "Countenance of Evaluation" (1967), he recommended that evaluators assess objectives and outcomes, but also antecedents and ongoing transactions (both intended and unintended). Stake's definition of evaluation, "an observed value compared to some standard" (Stufflebeam & Shinkfield, 2007, p. 414) would serve as an important influence on future teacher evaluation methods. By the mid 1970s, Stake had labelled his contributions to the evaluation to be "responsive

evaluation” because he surmised that goals or objectives may not be clear or known at the onset and therefore evaluators needed to respond given the actual circumstances. Many attributes of teacher evaluation reflect the components of responsive evaluation; for example, some of the methods used by responsive evaluators, such as expressive objectives and observation, are common methods used in teacher evaluation in the past and present (Stufflebeam & Shinkfield, 2007; Stake, 1967).

During this period, in the late 1960s and early 1970s, with the reconsideration and creation of approaches to evaluation, the evaluation of teacher effectiveness experienced new, impactful approaches. These approaches were used during this time to coincide with the changes in curricula and the renewed focus on educational quality (Hogan, 2007). A prominent approach developed in this time period was the method of *clinical supervision*. Robert Goldhammer’s book *Clinical Supervision: Special Methods for the Supervision of Teachers* (1969) provided a systematic approach to evaluating teachers through classroom observation that includes five stages:

1. Pre-observation conference
2. Classroom observation
3. Analysis (of the information obtained during the evaluation)
4. Supervision (post) conference
5. Analysis of the analysis (of the supervisor’s performance)

Cogan defined clinical supervision as “supervision focused upon the improvement of the classroom performance of the teacher by way of observation, analysis, and treatment of that performance” (Cogan, 1964, p. 118). Similar to medical “clinical” approaches,

clinical supervision collected data through observation in education, analyzed those data, and then provided steps to modify.

The Age of Realism reflects a major turning point in American educational history with the rapid involvement of the federal government in the funding and evaluation of school districts and educational programs. These years proved decisive in altering the course of evaluation theory and practice while also beginning to place greater emphasis on evaluations of teacher effectiveness.

The Age of Professionalism: 1972 to 2007

Evaluation as an academic field began to coalesce in the 1970s, evidenced through the creation of professional journals, the addition of evaluation methods courses to higher educational curricula, and the development of graduate programs in evaluation (Stufflebeam & Shinkfield, 2007; Hogan, 2007). The professionalization of evaluation as a field of study had broad implications for programs and personnel in a variety of fields, most especially education.

By the early 1980s, the education reform effort that was encompassing decades since the passage of the National Defense of Education Act was demonstrating that public investment and interest seemed to impact education quality little. The public began to increasingly describe the problem with education as a problem with teacher quality and ability. *Texas Monthly* published “Why Teachers Can’t Teach” (which went on to win the National Magazine Award for Public Service), which eviscerated the teaching profession by providing evidence that some teachers in Texas scored lower on standardized tests than average 16-year-old students. “Why Teachers Can’t Teach” provided an answer to the question posed and discussed: Teacher training focused too much on pedagogy and

not enough on subject matter mastery, although it was mentioned that teacher compensation in the sample area was probably too low to attract good candidates (Lyons, 1980). Inspired by ideas such as those in “Why Teachers Can’t Teach,” Secretary of Education Terrel Bell formed the National Commission on Excellence in Education with a mission to provide a “Marshall Plan” style solution to the problems in the American educational system (Goldstein, 2014). The resulting document, *A Nation at Risk: The Imperative for Educational Reform* refocused public and political attention on the lack of results to improve educational quality. The report’s introduction paralleled the national security focus and militaristic, Cold War language found in the creation and passage of the National Defense of Education Act in 1958:

If an unfriendly foreign power had attempted to impose on America the mediocre education performance that exists today, we might well have viewed it as an act of war. As it stands, we have allowed this to happen to ourselves . . . We have, in effect, been committing an act of unthinking, unilateral educational disarmament. (National Commission on Excellence in Education, 1958)

The report’s support for its termed “rising tide of mediocrity” included two decades of SAT score declines, non-academic high school course offerings, and a teaching force drawn from the lower levels of graduating classes. This evidence was offered in comparison to the results of more effective educational systems in other countries where economic product quality surpassed American output (specifically Germany and Japan). The report provided four areas of focus to remedy American education: better teachers, higher expectations for students (especially in high school), lengthening the school day

by one hour and the school year by 40 days, and more involvement and funding from the federal government. Amid a political climate of spending cuts and conservative concerns with federal involvement in state and local education curricula and decisions, policy-makers and the public focused on improving teachers as the road to improving American education (Goldstein, 2014; National Commission on Excellence in Education, n.d.)

Teachers' unions were initially mixed in their support for the recommendations in *A Nation at Risk*, but their resistance to evaluation systems that were considered subjective was a long-standing policy of the unions and its members. Classroom observation by administrators was considered subjective by the unions. The 1968 teacher strike in New York City was in response to an administrator's attempt to evaluate teachers in a subjective manner. Observations of teacher practice were commonplace in education by the early 1980s, but as Chester Finn stated, "The principals were often former gym teachers and had almost never been trained to be sophisticated overseers of teacher quality or performance. In the absence of quantitative data, a principal would be a perfunctory and ill-trained observer and either liked or didn't like what he saw" (Goldstein, 2014, p. 178). As Stufflebeam and Shinkfield stated, "Society has a critical need not only for competent evaluators but for evaluation-oriented decision makers as well" (Stufflebeam & Shinkfield, 2007, p. 117). What was clear from the tensions around classroom observations was that for teacher effectiveness reforms to be effective, reformers were going to have to focus on objective measures evaluated by competent supervisors.

A Nation at Risk prompted the Carnegie Forum on Education and Economy to publish the report *A Nation Prepared: Teachers for the 21st Century: The Report of the*

Task Force for Teaching as a Profession. This report, as the title suggests, seized on the recommendation in *A Nation at Risk* that teacher quality be improved. The report is responsible for introducing the concept of teacher accountability for student achievement. The teaching profession, according to the report, should be professionalized through accountability for student results, creation of a National Board of Professional Teaching Standards (to establish higher standards for teaching), improving higher education's teacher preparation programs, increasing teacher pay (Carnegie Forum on Education and the Economy, 1986). Today's teacher evaluation models and policies include provisions about teacher accountability that entered the education lexicon in the 1980s and 1990s. Reports like *A Nation at Risk* and *A Nation Prepared* were responsible for the establishment of teacher evaluation systems being put in place in 98% of school districts across the country.

The ability of the then developing field of evaluation to influence the creation of reliable and non-subjective evaluative models was vital to the teacher effectiveness reform movement. In the evaluation of personnel, including teacher effectiveness, an effort was undertaken by the Joint Committee on Standards for Educational Evaluation (1988) to create a set of standards for evaluating personnel. The resulting standards were designed with educational evaluation in mind but can serve as a framework for evaluating personnel in other fields as well. The result of this Joint Committee's work was the creation of The Personnel Evaluation Standards, which consists of four requirements:

1. The *propriety standards* require evaluations to be ethical and fair to the affected parties, including beneficiaries as well as the service provider.

2. The *utility standards* require evaluators to issue results that are credible, informative, timely, and influential. These results should help individuals and groups improve their performance and help superiors make needed personnel decisions and guide staff development and other personnel actions.
3. The *feasibility standards* require that evaluation procedures are efficient, politically viable, relevantly easy to implement, and adequately funded.
4. The *accuracy standards* require that evaluations provide sound information about a person's qualifications and performance. The results should be grounded in an up-to-date position description, take account of the particular work environment and institutional or societal mission, be based on systemic collection and analysis of data, and be validly interpreted and reported.

(Joint Committee on Standards for Educational Evaluation, 1988)

It is worth noting a number of merit pay movements existed in the 1980s education reform landscape. These methods, in theory, were welcomed by teachers and teachers' unions and reflected the influence of for-profit, business strategies influence on educational reform (as has been seen before, including with the Scientific Management movement 70 years earlier). However, in practice, these merit pay efforts eventually came to be opposed by teacher unions due to their inability to objectively evaluate teachers and their tendency to focus pay enhancements on only certain subsets of teachers. As the 1980s ended, the concept of teacher accountability remained prominent,

but the merit pay movement was all but extinguished, replaced by efforts centered on teacher collaboration and professional development (Brandt, 1995). Accountability would join with the standards-based movement of the 1990s to further shape the future of educational evaluation.

The 1990s brought the passage of new laws focused on improving instruction through teacher evaluation (Shough, 2010). Corresponding with the movement to create student standards, the teacher evaluation movement also sought to create standards of teaching. The formation of standards of teaching sought to provide a comprehensive view of teaching that would include what teachers should know and do to improve their teaching and student learning. These standards would allow evaluators to rate teachers in the different areas of teaching and create a comprehensive profile of the teacher's instructional abilities. Charlotte Danielson's *Enhancing Professional Practice: A Framework for Teaching* (1996, 2007) was a landmark publication within the field of teacher evaluation. Danielson's *Framework* (1997, 2007) became the first of many teacher evaluation models that would provide a standards-based approach to evaluating teaching.

The Framework, and other standards-based approaches to teacher evaluation, provided much more than the standard checklist-style approach to observing teacher practice. Evaluators could utilize *The Framework*, and others, to provide formative *and* summative data. Scriven's formative and summative evaluations had been growing in the evaluation literature and in usage by practitioners since the 1960s, but the combination of standards-based tools with the formative approach started to become the focus of teacher evaluation. Teacher evaluation as a formative experience could help in improving teacher

practice and in making decisions about professional development and not just termination decisions (Stufflebeam & Shinkfield, 2007). Danielson's and others' evaluation methods were well-researched in their creation and design, making these evaluation methods more "objective," which became important in beginning to gain the support of teachers' unions.

For many, the better researched and designed standards-based methods of evaluating teacher effectiveness was only part of the solution. The National Commission on Teaching and America's Future published *Doing What Matters Most: Investing in Quality Teaching* in 1997, in which the Commission advocated a comprehensive look at the teaching profession. The Commission called for reforms in standards, recruitment, preparation, professional development, and school restructuring to "move the country a giant step closer to meeting the goal of ensuring each student a qualified, competent, caring teacher by 2006" (Darling-Hammond, 1997, p. 39). The use of standards-based evaluation focused only on the teacher and did not factor in the effects of a teacher's instruction: student growth or achievement. By the 1990s, standardized testing was commonplace in schools; but with technological enhancements, testing could be more frequent, faster, and include greater levels of data analysis. Educational reformers, particularly in Texas, advocated using students' standardized testing results to hold teachers, schools, and districts accountable. Student achievement data began to be incorporated into teacher evaluation systems (Wiggins & McTigue, 1998). As the 1990s came to an end, decades of political and public concern, calls for greater accountability, standards-based assessment of both students and teachers, research from the field of

evaluation, and improvements in standardized testing technologies united to provide a clearer path to teacher evaluation than existed at any point in the 20th century.

The Age of Accountability: Teacher Evaluation in the 21st Century

Stufflebeam and Shinkfield created the periodization titles for The Pre-Tylerian Period, The Tylerian Period, The Age of Innocence, The Age of Realism, and The Age of Professionalism. They ended the titles in the “present,” which at the time of publishing was 2007) (Stufflebeam & Shinkfield, 2007). Peterson defined the period from 1960-1980 as the period of “accountability-based approaches” (Peterson, 1982, p. 45), but a review of the historical literature on teacher evaluation from the viewpoint of 2017 would better Title the early 21st century as “The Age of Accountability.” Stufflebeam and Shinkfield published in 2007, but The Age of Accountability is best defined by the passage of the No Child Left Behind legislation and continued through the Race to the Top grant program.

On January 8, 2002, President George W. Bush signed the No Child Left Behind Act (NCLB) into law. The sweeping update to the Elementary and Secondary Education Act sought to bring more requirements for districts in order to receive federal funds through greater accountability and targets (Klein, 2015). Specifically, with regard to teacher quality, NCLB specified that each state have *highly qualified teachers (HQT)* by 2005. The act defined HQTs as teachers having the following:

1. Bachelor’s degree
2. Full state certification or licensure
3. Proof that they know each subject that they teach (U.S. Department of Education, 2004)

This legislation codified for the entirety of the country what it meant to be a teacher.

Nationally defining a highly qualified teacher was major progress for educational reformers who were seeking to provide better teachers in each classroom.

Despite the increased accountability that NCLB brought to education, the law was somewhat disjointed in its efforts to bring comprehensive teacher evaluation due to its lack of language about the evaluation of teacher practice and teacher effectiveness.

NCLB focused largely on the aggregate school-level, and not individual teachers.

Schools that did not bring all students to proficient levels could be titled as “failing” schools and lose federal title funds. With regard to individuals, NCLB focused on what evaluation theory called “inputs” in personnel evaluation (in this case: teacher degree, years of service, certification, etc.) but stopped short of codifying requests to evaluate teacher practice (such as through standards-based evaluation models) and the linked student results (such as through students’ standardized test scores) (Rogers & Weems, 2010). Many argued that student achievement data must be a component of any comprehensive teacher evaluation policy in addition to teacher inputs and measures of teacher practice (Tucker & Stronge, 2005). Combining teacher qualifications with students’ standardized tests scores, even above the “individual teacher level,” raised the bar of expectations for schools, districts, and states while acclimating politicians, the public, and teachers (and their unions) to the use of student achievement data in evaluations.

Beginning in 2009, President Obama launched Race to the Top (RTTT). RTTT was a competitive grant designed to, as President Obama declared, “stop just talking about education reform and start actually doing it” (U.S. Department of Education, 2009).

According to the U.S. Department of Education, the grants were to be awarded to states that demonstrated commitments and plans in four educational reform areas:

- Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy
- Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction
- Recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most
- Turning around our lowest-achieving schools

(United States Department of Education, 2009)

All four reform areas would influence the adoption of new teacher evaluation systems in states and districts across the country. While the NCLB changes measuring schools, districts, and states were being enacted, researchers were working on methods to measure teacher-level effects on student achievement and those would be put into use to fulfill RTTT components (Goldstein, 2014). In many ways, RTTT served as a culmination of decades of teacher evaluation efforts and research by acting in three key domains of teacher evaluation: teacher quality (inputs), teacher practice (actions and abilities), and teacher effectiveness (student learning results). Since RTTT's creation in 2009 and the ESEA Flexibility Program of 2011, which provided waivers from NCLB targets for state's adoption of RTTT elements, 40 states now require some measures of student academic growth in teacher evaluations (Walsh, Joseph, Lakis, & Lubell, 2017).

Teacher accountability arrived at the beginning of the 21st century; and with the recent reauthorization of the Elementary and Secondary Education Act, titled Every Student Succeeds Act (ESSA), teacher accountability will persist. ESSA's impact on teacher evaluation is found in four of its priorities:

1. Annual statewide assessments of all student learning
2. Student performance targets and school ratings
3. Accountability, interventions, and supports for struggling schools
4. Competitive program to evaluate and reward effective educators in high need schools (U.S. Department of Education, n.d.)

For decades throughout the 20th century, global and national issues thrust teachers to the center of the debate over the strength of the American economy and democracy. Trillions of dollars were invested in education, much targeted at helping to close the achievement gap and improve student outcomes. Much was debated about teacher evaluation; but compared with the early 21st century's impact of NCLB and RTTT, there was little change over those early decades. The 21st century saw the combination of the standards-based evaluation of teacher practice, student achievement results, and a public that wants to see both functioning effectively for all teachers.

The Framework for Teaching

Throughout much of the 20th century, a clear definition of “good teaching” could not be attained. Charlotte Danielson worked to develop her *Framework* to create a common understanding of good teaching, as she stated in the following:

A framework for professional practice is not unique to education.

Indeed, other professions – medicine, accounting, and architecture,

among many others – have well-established definitions of expertise and procedures to certify novice and advanced practitioners. Such procedures are the public’s guarantee that the members of a profession hold themselves and their colleagues to high standards of practice (1996, 2007, p. 2).

Danielson’s justification for the creation of the *Framework* echoes the decades of debate from within and without the educational community. Decades of attempts at teacher effectiveness tools demonstrated that teaching is complex, and that complexity led to disagreement over evaluating a teachers’ effectiveness; but Danielson’s *Framework* provided a “shared understanding of teaching” (Danielson, 1996, 2007, p. 2). The shared understanding would become useful in the development of teachers’ skills and also in the evaluation of their effectiveness.

The *Framework* is a standards-based evaluation of teacher performance. Since measures of a teacher’s skills are collected through data (including through observation, conversation, or portfolio) of a teacher’s actions, it could also be termed a performance-based evaluation. Danielson posited that some major reasons for the creation of the *Framework* were to reflect the complexity of teaching, to provide for a common language for professional conversation, and to provide a structure for self-assessment and reflection on practice (Danielson, 1996, 2007). Indeed, states and districts that have chosen Danielson’s *Framework* as the evaluation model for evaluating teachers heavily rely upon the ability of the *Framework* to provide meaningful conversation between supervisors and teachers and provide for reflection in the process of improving teaching practices. Danielson articulated that the *Framework* exists due to several assumptions:

important learning for students, the nature of learning and how to promote it, the purposeful nature of teaching, and the nature of professionalism.

In formulating the *Framework*, Danielson built upon the vast body of research on teaching and teacher effectiveness that began in earnest in the 1960s. Specifically, beginning with the teacher effects research of the 1960s and 1970s and continuing with the work of Hunter in the 1980s, the importance of instruction on student learning was becoming better defined. Danielson does acknowledge that much of the teacher effects and Hunter research proved to be ineffective in consistently supporting positive student learning. The teacher-centered, structured classrooms that resulted from the teacher effects and Hunter research are only a small portion of what the *Framework* constitutes as effective teaching, albeit they were important contributions to the research literature on teaching. A major shift in the research on teaching occurred in the 1980s and 1990s with a movement away from teaching through the perspective of behaviorism to cognitive learning theory (Danielson & McGreal, 2000). Changing demands of the economy and society in these decades caused the need to provide for students with different sets of skills that included problem solving, life-long learning, critical thinking, and the like. Cognitive learning theory provided much of the literature that led to the development, or perhaps revival, of constructivism (the approach to learning that learners, both adult and child, learn by constructing their understanding of what they study) (Eggen & Kauchak, 1996). Dewey, Vygotsky, and Piaget laid the groundwork and research for constructivism, but that view of learning had fallen out of fashion within American education after the National Defense of Education Act refocused the country on “basic skill” development of students in 1958. With the revival of this theory of learning

through cognitive learning theory, Danielson approached the creation of the *Framework* from the viewpoint of constructivism in teaching and learning, acknowledging the role teachers must play in guiding the process but stating, “A teacher using a constructivist approach recognizes that if students are to understand the concept, they must do much of the intellectual work themselves” (Danielson, 1996, 2007, p. 16). Danielson drew on the vast amount of literature on teaching and psychology to arrive at the formation of the *Framework* as a set of standards that are empirically and theoretically supported. As she and McGreal stated, “Effective evaluation and professional development programs must start with a rich set of teaching standards that reflect what we know” (Danielson & McGreal, 2000, p. 15).

Grounded in empirical and theoretical research on excellent teaching, the *Framework* is comprehensive, generic to all disciplines, publicly available for reviewing, independent of a particular teaching methodology, and coherent in its structure. The structure consists of four domains, 22 components, and 76 elements. The domains encompass the four macro areas of teaching: Domain 1, Planning and Preparation; Domain 2, The Classroom Environment; Domain 3, Instruction; and Domain 4, Professional Responsibilities. The four domains are of relatively equal weight with five or six components each and include two domains (Domains 1 and 4) that are primarily practices teachers would engage in outside of direct classroom instruction with students and two domains (Domains 2 and 3) that are practices teachers would engage in while in direct classroom instruction with students. The components include Domain 1: demonstrating knowledge of content and pedagogy, demonstrating knowledge of students, setting instructional outcomes, demonstrating knowledge of resources,

designing coherent instruction, and designing student assessments; Domain 2: creating an environment of respect and rapport, establishing a culture for learning, managing classroom procedures, managing student behavior, and organizing physical space; Domain 3: communicating with students, using questioning and discussion techniques, engaging students in learning, using assessment in instruction, and demonstrating flexibility and responsiveness; Domain 4: reflecting on teaching, maintaining accurate records, communicating with families, participating in a professional community, growing and developing professionally, and showing professionalism. This structure, along with the elements of the *Framework* provide a comprehensive, balanced approach to understanding the complexities and necessities of teaching.

Danielson's *Framework* builds off decades of teacher evaluation research, but the change in approach is notable from past evaluation methods that were focused on the supervision of teacher "behaviors." Clinical supervision, for example, held that teaching is complex behavior and as such can be broken down into elemental behaviors that can be systematically developed in the teacher by training techniques" (Weller, 1971, pp. 12-13). Danielson's *Framework* refutes the position that the correct "elemental behaviors" can improve instruction, while underscoring that teaching is complex and multifaceted.

Theoretical Frameworks

Michael Scriven

Michael Scriven's research and writing is paramount to understanding the goals and purposes of evaluation. Stufflebeam and Shinkfield (2007) termed Scriven's theory of evaluation as "consumer-oriented evaluation," in part due to Scriven's 1969 statement that the evaluator is "an enlightened surrogate consumer." In the field of education, the

“consumers” would be students. To Scriven, evaluators must go beyond just measuring things and deciding if goals were met, but creating defensible value judgments as well. Scriven (1974) introduced the evaluation concept of “goal-free evaluation” where “regardless of goals, the evaluator must identify outcomes and assess their value from the perspective of consumers’ needs” (Stufflebeam & Shinkfield, 2007, p. 369).

In “The Methodology of Evaluation,” Scriven introduced the terms “formative evaluation” and “summative evaluation” (Scriven, 1967). Scriven believed that the goal and role of evaluation were not always properly understood by evaluators. To Scriven, the goal of evaluation was to judge the value, but the roles could vary widely and the confusion between the goal and the many roles would often lead to poor evaluations. Scriven’s solution to the confusion was the creation of two main roles for evaluation: formative and summative. He later added “ascriptive evaluations,” but those are not relevant to the discussion of teacher evaluation models (Scriven, 1967; Stufflebeam & Shinkfield, 2007). Scriven saw formative and summative as working somewhat together. As he described, the formative evaluation would occur during the development of the product or service, and the summative would occur after the development was completed. The final service or product, however, would have been formed, in part, by the information from the formative evaluation stage(s). Formative and summative are now considered standard terminology in the evaluation field (especially in education). In 1967, Scriven gave educational administrators these keystone terms: “formative” to help improve and “summative” to help compare (Scriven, 1967; Stufflebeam & Shinkfield, 2007).

Central to the current debate over teacher evaluation has been teacher attributes (degree, years of experience, etc.) versus student achievement. Scriven first identified these forms of evaluation as “intrinsic evaluation” and “payoff evaluation”. Intrinsic evaluation is unconcerned with results, whereas payoff evaluation is only concerned with results. For decades, teacher evaluation has been dominated by intrinsic evaluation, however in the age of Race to the Top there has been a shift towards a more balanced approach that includes student achievement outcomes in the evaluation of teachers (Stufflebeam & Shinkfield, 2007).

When it comes to the final stages of evaluation, the synthesis of data to form a judgment, Scriven is unwavering that what is needed “are clear – but not simplistic – rules for deciding whether and how to reach justified conclusions” (Stufflebeam & Shinkfield, 2007, p. 381). If a rule is impossible, Scriven encourages evaluators to take other steps, including creating rubrics and/or calibrating evaluators. According to Scriven, the goal would be either through rule, rubric, calibration, or some other method that different evaluators who are evaluating the same data would come to the same conclusion. If this is impossible, Scriven contends that a final judgment should not be offered; instead, performance levels for separate aspects of the evaluation should be provided (Scriven, 1994).

Fullan Leading Learning

A learning leader differs from an “instructional leader” because instructional leaders tend to be more micro-focused on working with teachers to improve individual instruction. Fullan (2014) argues that this micro-focus, which began in the late 1980s, is not beneficial to driving improvements in instruction, and is often impossible given the

size of schools and the non-instructional demands of today's principalship. Principals' concerns with instructional involvement are not new; Fullan cited a 1984 survey of principals in Ontario, Canada, who stated that the most drastic changes to their roles in those years were "teacher performance review and curriculum implementation." Fullan's learning leader (or "leading learner," or "lead learner") is "one who models learning but also shapes the conditions for all to learn on a continuous basis" (Fullan, 2014, p. 9). The literature on effective principals mostly coalesces around the themes of collaborative cultures, learning communities, and capacity building (DuFour & Marzano, 2009; DuFour & Fullan, 2013; Fullan, 2010; Leithwood & Seashore-Louis, 2011; Hargreaves & Fullan, 2012). These actions are the "right drivers" in contrast to what Fullan argues are the "wrong drivers" (Fullan M, 2014).

The right versus wrong drivers that are related to teacher evaluation include accountability (wrong) vs. capacity building (right) and individualistic solutions (wrong) vs. collaborative effort (right) (Fullan, 2014). Accountability measures that are, largely, tied to "carrot and stick" methods may be effective for some shorter-term, less complicated work, but they are not effective motivators for long-term, professional work (such as teaching) (Pink, 2009). Teacher reform efforts have been, largely, targeted at improving individual teachers, but Fullan argues that more success in providing highly effective teachers has come from systems that "help teachers work together in a focused way to use diagnostic student data linked to the improvement of instruction in order to get better results" (Fullan, 2014, p. 31). It is through this process that professional capital is built.

Fullan provides an example of professional capital in action, as stated in the following:

Leading learners are principals who ensure that the group focuses on a small number of key elements: specific goals for students; data that enable clear diagnosis of individual learning needs; instructional practices that address those learning needs; and teachers learning from each other, monitoring overall progress, and making adjustments accordingly. All of this is carried out in a developmental climate (as distinct from a judgmental one) with norms of transparency within and external to the school. Within this set of conditions, accountability measures, including teacher evaluation, can and do occur, but they are conducted within a culture of collaborative improvement. (p. 63)

With regard to formal teacher evaluation processes and the process of feedback associated with them, Fullan argues that there is little evidence to conclude that simply utilizing an evaluation system to root-out poor teachers will be effective. Instead, he provides evidence that the teacher evaluation systems work but only when integrated into group work. As Fullan sums-up, “You can use both—teacher evaluation and collaborative cultures—just get the order right” (Fullan, 2014, p. 77). Teacher evaluation models, such as Danielson’s, are important for providing a framework for collaborative discussion, sharing, and learning among colleagues. It is with tools like the professional teaching standards of the Danielson *Framework* that Fullan’s leading learners can create

the conditions for teachers to learn both with the principal and with teacher colleagues and therefore build professional capital.

Summary

The evolution of the field of evaluation and education are intertwined, and Chapter II provided a review of the existing literature pertaining to teacher evaluation and principals' perception of the effectiveness of the evaluation process within the context of the evolution of the field of evaluation. The theoretical frameworks of Michael Fullan and Michael Scriven provide the lens for the coming analysis. The literature review concluded with an analysis of Charlotte Danielson's *Framework for Teaching*, which provides a comprehensive view of evaluation within this model.

CHAPTER III

METHODOLOGY

Introduction

The purpose of this study was to examine principals' perceptions of their ability to effectively use Danielson's *The Framework for Teaching* to provide formative feedback to teachers, make summative decisions on tenure and termination, and improve their instructional leadership of the school. Since I was looking to discover principals' perceptions, a phenomenological, qualitative study was most appropriate. This chapter will provide the design and structure of the methodology for this study. The chapter begins with background information about myself in terms of my professional experiences, continues with the design of the study, profiling of the research site, and information on participant sampling. The chapter then provides the process for data collection, data analysis, and protection of human subjects and ethical considerations. Finally, the chapter concludes with sections on validity and reliability and the role of the researcher and research bias.

Background

I began teaching in the 2003-2004 school year in a non-public, Catholic school. I was at the high school level, in which I had five direct supervisors (four of whom co-supervised instruction). At that time, I was a member of a large teachers' union that was responsible for negotiating the terms for teacher evaluation in all member schools (approximately 175 schools). In my first school year, I was observed twice. Both observations were conducted formally; that is, with notice. Both of those observations, separately conducted by my department chair and my principal, included a pre-

conference, the observation, a post-conference, and then a written observation report to follow. At both pre-conference meetings, I was asked to provide the topic of the lesson and then verbally walk the observer through some of the procedures and materials to be used. During the conference, the observer used a checklist-style observation form. The observation form consisted of four sections (Instructional Process, Effective Planning, Pupil-Teacher Relationship, and Classroom Environment) with indicators that were not balanced in their number across the sections. Each of the 30 indicators was rated using the following ratings: Good, Satisfactory, Needs Improvement, Unsatisfactory, and Not Applicable. The observer was only required to provide comments for Needs Improvement or Unsatisfactory ratings. Those comments for Needs Improvement and Unsatisfactory needed to contain constructive suggestions for improvement. Despite that I had only been teaching for a matter of months and that I had no prior training or college instruction in education, I did very well. I received absolutely no ratings of Needs Improvement or Unsatisfactory. I received some suggestions to improve my behaviors; for example, waiting longer after asking a question and where to write notes on the board. Now looking from the perspective of teacher evaluation in 2017, it is incomprehensible that in not a single area did I “Need Improvement” in the complex profession of teaching. However, this rating was to be expected because nearly all teachers did well on these observations. If I wanted to be successful in this system, I would simply need to address these relatively minor teacher behavior issues for the next formal observation because in my first year no supervisor ever observed, even briefly or informally, my lessons outside these two formal observations.

As I, and other teachers in the system, moved from day to day and year to year, we lacked a high quality and generally accepted standard of teaching quality. We lacked the ability to see our true strengths and our areas in need of improvement. We lacked the specific feedback from our supervisors that was necessary to help us develop as professionals. Not surprisingly, I came to view observations as perfunctory, where nothing was to be gained and there was little chance anything would be lost. The collective bargaining agreement included an Annual Professional Performance Appraisal. However, in my eight years as a teacher, only in the final three years was this required by the principal; and in those three years it was used only as a teacher's "self-evaluation." The annual appraisal consisted of three categories with indicators for each, but these categories and indicators were only loosely correlated with the formal observation form. Three years in, after five total formal observations, two walkthroughs observations, and zero annual reviews, I was granted tenure.

After 10 years in education, I became a principal in September of 2013. I changed schools but remained in the same system and now encountered the teacher evaluation process from a new perspective. At first glance, and from the perspective of 2017, it would make sense to lay blame on my previous administration that was ineffective in providing the necessary feedback for growth for myself and other teachers. However, the observation tool that was being used asked observers to look for specifically defined teacher behaviors or student behaviors but did not include a description of what those behaviors should look like. Additionally, because there was no rubric for deciding on ratings, the observers had only their own thoughts and experiences to decide where particular behaviors belonged on the rating scale. The observation form did show

thoughtful intention by the parties of the collective bargaining agreement. The indicators on which I was rated sampled pieces of different movements in teacher evaluation throughout the 20th century. For example, “Exhibits poise, voice control, and tact” hearkens back to the era when teacher character traits informed teacher quality; “Provides enrichment” followed the checklist-oriented teacher behaviors movement; and “Prepares assessment tools and techniques based on learning standards” showed the influence of Tyler’s objective/outcome. There was certainly effort to provide for good instruction in the classroom; but as I was now experiencing, the principals and other instructional supervisors lacked the proper tools that a comprehensive, standards-based teacher evaluation model would provide. Without the proper tools, the purpose for spending time with teachers in their classrooms and in discussion about their instruction seemed a burden and fruitless.

Two years into my principalship and one year prior to beginning my research, my school system made a dramatic and sudden shift to the Danielson *Framework* for both teacher classroom observation and as an annual performance review. As principals, we were now plunged into the world of instructional leaders and evaluators in a way that did not exist before. In addition to utilizing the *Framework* as a comprehensive approach, the annual quantity of evaluations required to meet the collective bargaining minimums, at the least, doubled for most principals. Evaluating teachers went from perfunctory to thorough almost overnight.

By the time I was prepared to begin my research in 2016, the ground had shifted, both nationally and within my school system, in teacher evaluation. While our school system was not utilizing student test scores as components of the approach, many

neighboring districts and states were (as required by the RTTT grants). Almost universally, public and Catholic school systems had adopted comprehensive, standards-based teacher evaluation models. I struggled with how best to work with teachers in this new system and now with a more professional goal: to help teachers improve their teaching and therefore their effectiveness through meaningful evaluation. How was I to be an effective instructional leader, coach, and final evaluator for the teachers? Coming from an experience in which I had received little feedback about my teaching, how should I provide feedback to teachers? Of course, the purpose of the Danielson *Framework* is to improve teaching and teacher effectiveness, but this improvement does not happen in a vacuum. Principals play a crucial role as formative and summative evaluators, and research needs to be done to inform practices for principals in effective methods for evaluating teachers and thereby improving teacher effectiveness.

My experiences and a review of the teacher evaluation literature assisted in the formation of the following five research questions that guided this study:

1. How do principals perceive their preparation to utilize the Danielson *Framework*?
2. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is effective in providing formative feedback to teachers?
3. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is beneficial to the tenure and termination processes?
4. To what extent, if any, do principals perceive that the use of the Danielson *Framework* has improved their instructional leadership?

Design

I utilized the qualitative method for this research design. Merriam (2009) defined qualitative research as research “interested in *understanding the meaning people have constructed*; that is, how people make sense of their world and the experiences they have in the world” (p. 13). Qualitative research embodies the following characteristics: focus on meaning and understanding; researcher as primary instrument; an inductive process; rich description; and usually emergent, flexible, purposeful sampling (Merriam, 2009).

Merriam (2009) and Creswell (2009) explained that qualitative research comes from a perspective that no experience, or phenomenon, has a singular reality and that all realities are subjective because they are constructed by individuals with different pasts and experiences constructed within a social setting. Qualitative phenomenography is the study of how people experience and describe a phenomenon deriving from the perspective “that we can only know what we experience by attending to perceptions and meanings” (Patton, 2002, p. 105). Since the goal of this study was to understand the perspectives of principals using the Danielson *Framework* a phenomenological approach was best suited.

In seeking to make meaning of principals’ perceptions, this method needed to be flexible enough to allow for emergent themes to evolve throughout the study. Unlike with quantitative research, qualitative research allows for the data analysis to be ongoing throughout the data collection period rather than after data collection. This permits the researcher to begin constructing themes and patterns from earlier in the process.

Qualitative research is contextual, and what participants said were their feelings and perceptions about teacher evaluation was considered the reality. Through semi-structured

interviews I uncovered principals' feelings and beliefs that then helped to uncover their perceptions about the impact on their instructional leadership, if any, of using the Danielson *Framework* for teacher evaluation.

Profile of the Site

For the purpose of this study a pseudonym, Seven Fountains, was used in place of the actual school district. Seven Fountains was chosen for this study because this non-public school district chose Charlotte Danielson's *Framework for Teaching* as its teacher evaluation model beginning in the Fall 2015-2016 school year. By this time within New York State, public school districts were mandated to utilize a state-approved teacher evaluation model. Following the public model, many non-public schools, including all schools within Seven Fountains, also transitioned to new evaluation models. Seven Fountains is one non-public district within a larger grouping of nine non-public districts. Each district has its own superintendent and is structured under an overall superintendent of all nine districts. The districts, such as Seven Fountains, title themselves "regions" but function as community school districts in many bureaucratic ways, including instructional, financial, operational, and personnel areas. For the purpose of this study to utilize the common vocabulary in education, I refer to Seven Fountains as a district. Seven Fountains consists of 25 PreK-8 schools with a total enrollment of over 6,500 students. The table below reflects Seven Fountains' 2016 percent proficient for Grades 3-8 ELA and Math exams.

Table 1

Seven Fountains' 2016 New York State Exam Percent Proficient

Seven Fountains' 2016 New York State Exam Percent Proficient		
Grade	English Language Arts (ELA)	Mathematics
3	54%	49%
4	51%	52%
5	44%	42%
6	48%	42%
7	48%	39%
8	53%	36%

Note. Information accessed from Seven Fountains district website in 2017.

The Seven Fountains' teacher evaluation process is generated through collective bargaining with the teachers' union and outlined in the Collective Bargaining Agreement (CBA). Going into the 2015-2016 school year and for more than 15 years, the teacher evaluation process remained unchanged in Seven Fountains. The former evaluation process utilized a checklist-style classroom observation form and an annual review form that was quite dissimilar from the observation form. The CBA required no mandatory observations except some (1-2) for non-tenured teachers, and principals had complete discretion to offer tenure to teachers after their third consecutive year of teaching at a school. With the enactment of a new CBA and teacher evaluation model in the fall of 2015, principals were faced with a dramatic shift in the process. Requirements in the number of observations increased (as outlined in the chart below), the new evaluation would be based on the Danielson *Framework*, and the annual performance review was

directly correlated with the *Framework* (although just district policy and not in the CBA). Principals needed to provide evidence to the superintendent in recommending tenure for a teacher. The new process was designed to both improve instruction for all teachers and provide a more reliable method for summative decisions (termination and tenure).

Table 2

Seven Fountains' Observation Number Requirements

Seven Fountains' Observation Number Requirements (Beginning with new CBA in 2015-2016 school year)		
Teaching Year/ Status	Number of Observations	Type of Observations
1 st Year	4	<ul style="list-style-type: none"> - Announced within the first 45 days - Announced within the first 90 days - Unannounced within the first 90 days - Announced and after the first 90 days
2 nd Year	3	<ul style="list-style-type: none"> - Announced within the first semester - Unannounced and within the first semester - Announced and within the second semester
3 rd Year	3	<ul style="list-style-type: none"> - Announced within the first semester - Unannounced and within the first semester - Announced and within the second semester
Non-tenured in the 4 th Year or more	3	<ul style="list-style-type: none"> - Announced within the first semester - Unannounced and within the first semester - Announced and within the second semester
Tenured	2	<ul style="list-style-type: none"> - Announced and within the first semester - Unannounced and within the first or second semester

Sampling

Seven Fountains provided a total population of 25 principals. Principals with at least one complete year of experience as a school building leader in the district were invited to participate in the research by way of a research recruitment letter sent via email to each of the principals employed in the Seven Fountains district. The email was approved by the superintendent of the district. The 12 principals 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 interview. Selection of the participants from among the volunteers was purposeful. It is important to explore the perceptions of principals with a range of experience to determine if that experience influences their perspective of the influence of the Danielson *Framework* on their leadership. Glesne (1999) suggested using “criteria that the literature and your experience suggest are particularly important” (p. 30). The criteria used for inclusion in this study were the following:

- Gender representation
- Representation of different ages and years of instructional leadership/coaching experience
- Minimum of one year of experience as a principal within the school district

All participants were issued pseudonyms (a participant letter) to protect their anonymity, as with the school district. Volunteers who were not selected for participation in this study received a letter of thanks for their interest in participating.

Table 3

Summary of Demographic Profile Questionnaire for Each Participant

Summary of Demographic Profile Questionnaire for Each Participant				
Principal	Sex	Number of Years as a Teacher	Number of Years of Instructional Leadership/ Coaching Prior to Principalship	Number of Years as a Principal
A	Male	6-10	6-10	6-10
B	Female	15+	1-5	1-5
C	Male	6-10	6-10	1-5
D	Female	6-10	15+	15+
E	Female	15+	6-10	1-5
F	Male	1-5	6-10	6-10
G	Female	15+	6-10	1-5
H	Female	15+	15+	1-5
I	Female	1-5	11-15	11-15
J	Female	11-15	6-10	1-5
K	Female	6-10	6-10	6-10
L	Female	6-10	15+	15+

Data Collection

Because the purpose of this study was to reveal the perceptions of the Danielson *Framework's* impact on the instructional leadership of principals, I collected qualitative

data in the form of interviews with principals and school documents related to their evaluation practices. Each of the 12 principals were interviewed through a semi-structured, in-person, 35-45 minute interview and were recorded using a recording device. The interviews were conducted privately, and I transcribed all interviews within one week of the session. The interviews helped me to learn about the principals' perceptions of the impact of the Danielson *Framework* on their instructional supervision and overall teacher evaluation. The documents and artifacts helped me to identify the evaluation in practice and reveal the nature of the application of the teacher evaluation model.

Table 4 provides an overview of the data collection process. Research Question 1 and its correlating sub-questions are identified and described below as an example.

Table 4

Research Questions and Corresponding Interview Questions

Research Questions	Corresponding Interview Questions
1. How do principals perceive their preparation to use the Danielson <i>Framework</i> ?	<p>1a) Do you feel you have the necessary training and practice to evaluate teacher performance accurately in all domains?</p> <p>1b) Do you feel your understanding of the differences between the HEDI ratings is aligned with that of your principal colleagues?</p> <p>1c) To what extent, if any, have you participated in calibration observations with fellow administrators?</p> <p>1d) To what extent, if any, do you feel the Danielson <i>Framework</i> evaluation model differs from your experience with evaluation as a teacher?</p>
2. To what extent, if any, do principals perceive that the use of	2a) How, if at all, has the Danielson <i>Framework</i> helped you provide formative feedback to teachers?

the Danielson <i>Framework</i> is effective in providing formative feedback to teachers?	<p>2b) How, in using the Danielson <i>Framework</i>, has this feedback impacted your relationship with teachers?</p> <p>2c) How would you describe the professional conversations between you and your teachers since using the Danielson <i>Framework</i>?</p> <p>2d) How, if at all, has the Danielson <i>Framework</i> improved the quality of professional conversations between you and your teachers?</p>
3. To what extent, if any, do principals perceive that the use of the Danielson <i>Framework</i> is beneficial to the tenure and termination process?	<p>3a) What impact, if any, has the use of the Danielson <i>Framework</i> had on your role as a coach?</p> <p>3b) What impact, if any, has the use of the Danielson <i>Framework</i> had on your role as an evaluator?</p> <p>3c) How, if at all, has the use of the Danielson <i>Framework</i> helped establish expectations for recommendation for tenure?</p>
4. To what extent, if any, do principals perceive that the use of the Danielson <i>Framework</i> has improved their instructional leadership?	<p>4a) How, if at all, has the Danielson <i>Framework</i> helped to define your expectations for teachers' classroom performance?</p> <p>4b) How have you used the Danielson <i>Framework</i>, if at all, when you design and implement professional development?</p> <p>4c) In what ways, if any, do you think your approach to teacher evaluation has been influenced by the use of the Danielson <i>Framework</i>?</p>

Table 5 provides some sample interview questions derived from the theoretical frameworks of the study.

Table 5

Sample Interview Questions Derived from the Theoretical Frameworks

Sample Interview Questions Derived from the Theoretical Frameworks		
Sample Questions	Theoretical Framework	Theorists
How, if at all, has the Danielson <i>Framework</i> helped you provide formative feedback to teachers?	Formative Evaluation	Scriven
How would you describe the professional conversations between you and your teachers since using the Danielson <i>Framework</i> ?	Leading Learner	Fullan
How have you used the Danielson <i>Framework</i> , if at all, when you design and implement professional development?	Leading Learner	Fullan
	Formative Evaluation	Scriven
What impact, if any, has the use of the Danielson <i>Framework</i> had on your role as a coach?	Lead Learner	Fullan
	Formative Evaluation	Scriven
How, if at all, has the use of the Danielson <i>Framework</i> helped establish expectations for recommendation for tenure?	Summative Evaluation	Scriven

Approval to conduct this research was granted by the superintendent of Seven Fountains (See Appendix A). Once I received approval from the Seton Hall University Institutional Review Board (IRB) (See Appendix B), I began to conduct the interviews within the district sites. Before conducting the interviews, I formed an expert panel that consisted of three experienced principals to field test the questions. The panel was organized and assembled to provide feedback on the structure and quality of the

questions in terms of their clarity and transparency. The feedback received from the panel was used to revise the questions as needed prior to conducting the interviews. No members of the expert panel participated in the study, and all were from outside the Seven Fountains district. The interview questions were created specifically for this study based on a review of the literature and the conceptual theories for the study. The questions were designed to gain the most data regarding principals' perceptions of the Danielson *Framework's* ability to enhance their supervision of instruction and teacher evaluation.

The interviews were conducted in a semi-structured format to allow for a deeper understanding of the participants' responses through follow-up questions. Semi-structured interviews permit for a consistency between interviews and a structure to ensure that the focus data are collected but allow for the researcher to use follow-up questions to seek clarification. The semi-structured approach results in interviews that vary considerably, while still providing for the data necessary to address the purposes of the study. The semi-structured approach is highly useful when seeking to reveal perceptions and helps the researcher find both commonalities and differences among participants' responses (Miles & Gilbert, 2005).

Field notes were taken during each interview. The field notes allowed me to record visual aspects of the interview while listening and recording audio on a recording device. Visual information recorded during the interviews included body movements, hand gestures, eye movements, etc. The use of field notes can be helpful for the researcher during analysis by providing the research with richer data about the interview encounter (Spencer, Ritchie, & O'Connor, n.d.). At the conclusion of each interview, I

created reflective notes, which Bogdan and Biklen (2003) described as useful for identifying a researcher's feelings as related to each individual interview. By reviewing field notes, interview recordings, and reflective notes, I was more accurately able to interpret participants' responses from the triangulation of interview data sources.

At the end of each interview I used member checking to ensure participants were able to review responses and make any revisions or additions to those data. After I transcribed each interview, I emailed the transcription to the participant for his or her review. This allowed each participant to make edits and additions to the interview or to make corrections to any error I may have committed in the transcription process. Through this member checking process of review and revision, the validity of the data and the study is stronger (Creswell & Miller, 2000).

All data and documents kept for the study are kept in a locked file cabinet. The documents include field notes, flash drives, transcripts, questionnaires, and other printed papers. To protect the identity of the participants, all the audio recordings were destroyed after the transcript was approved by the participant. All data will be kept for a period of three years, and then it will be destroyed.

Data Analysis

I documented demographic descriptors for the purpose of exploring patterns or themes during my data analysis. The following are the descriptors that were documented and coded for each participant: range of years as a teacher (range used to protect anonymity), range of years of instructional leadership prior to principalship (range used to protect anonymity), range of years as a principal (range used to protect anonymity), and sex.

Through interviews and thoughtful examination of teacher evaluation artifacts, I sought to understand the principals' perceptions of the effect the Danielson *Framework* has had, if any, on their instructional leadership. The research questions and theoretical frameworks provided the focus and direction of these analyses. The data were coded. After reading all materials, drafting initial codes, and creating finalized codes, I was able to identify themes.

Table 6 provides a preliminary listing of the data-driven codes.

Table 6

Preliminary Set of Data-Driven Codes

Preliminary Set of Data-driven Codes			
Authenticity	Accuracy	Capacity Building	Calibration
Clear Expectations	Coach	Collaborative Cultures	Compassion
Comprehensive	Conversation Starter	Difference of Experience	Feasibility
Focus	Formative Feedback	High Expectations	Instructional Leadership
Leading Learning	Learning Communities	Preparation	Principal Colleague Support
Professional Development	Propriety	Relationships	Self-taught

Specificity	Standards for Education Evaluation	Summative Feedback	Teacher Personality
-------------	--	-----------------------	---------------------

Table 7 Provides emerged themes as related to the research questions.

Table 7

Emerged Themes as Related to the Research Questions

Emerged Themes as Related to the Research Questions	
Research Question	Themes
1. How do principals perceive their preparation to use the Danielson <i>Framework</i> ?	<ul style="list-style-type: none"> - Differences from past experiences - Self-directed preparation and training - Variability in ratings and relevance
2. To what extent, if any, do principals perceive that the use of the Danielson <i>Framework</i> is effective in providing formative feedback to teachers?	<ul style="list-style-type: none"> - Useful tool in providing formative feedback and coaching - Focused conversations about teacher practices - Relationships and personalities
3. To what extent, if any, do principals perceive that the use of the Danielson <i>Framework</i> is beneficial to the tenure and termination process?	<ul style="list-style-type: none"> - Clear expectations for job performance - Evidence to determine teacher tenure
4. To what extent, if any, do principals perceive that the use of the Danielson <i>Framework</i> has improved their instructional leadership?	<ul style="list-style-type: none"> - Improving instructional leadership practices - Setting high expectations - Principals as leading learning

Protection of Human Subjects and Ethical Consideration

Throughout this study, ethical research practices were consistently applied. I completed the National Institute of Health's course "Protecting Human Research

Participants” (Appendix C). The American Psychological Association identifies five Guiding Principles in its Ethics Code. These five Guiding Principles were utilized by me throughout the study and are the following:

- Principle A: Beneficence and Nonmaleficence
- Principle B: Fidelity and Responsibility
- Principle C: Integrity
- Principle D: Justice
- Principle E: Respect for People’s Rights and Dignity

(American Psychological Association, 2017)

Participation in the study was voluntary; all participants completed an informed consent form, explaining the study’s purpose and any risks and benefits of participation.

Participants’ anonymity was always maintained. As described earlier, each participant was able to review the transcript of his or her interview for edits and additions prior to its use for analysis.

Validity and Reliability

In both quantitative and qualitative research, validity is essential to ensure the findings have merit and worth. Unlike in quantitative research, in qualitative research the main tool is the researcher. Paramount to ensuring validity in qualitative research is the trustworthiness of the researcher to accurately understand the meanings of the participants. Without an accurate reporting and analysis of the participants’ meanings, the study and its results would be flawed (Golafshani, 2003; McMillan & Schumacher, 1997).

In order to secure validity in this study, I included the following actions advocated by Creswell (2002a, 2000b) and Roberts, Priest, & Traynor (2006):

- Use of an expert panel to review interview questions prior to the interview collection period. These panel members were able to help bring clarification to the interview questions.
- Interviews were transcribed with the use of an audio recording from the live interview session. The creation of a transcript with the aid of an audio recording provided for better accuracy of diction and meaning.
- Field notes were taken during the interviews. The field notes allowed me to have a greater understanding of the meaning of the participants' responses by noting the expressions, body movements, and tone of the participants during different points of the interview.
- Including a stage of member checking allowed the participants to correct their interview transcripts so that they may represent their feelings and meanings in the most accurate way.
- Utilizing qualitative analysis software, Dedoose, and a priori coding developed through a thorough review of the literature ensured greater reliability in the analysis of the data collected.

This study researched the perceived influence, or lack thereof, of the Danielson *Framework for Teaching* on a principal's ability to be an effective instructional leader. Given that the research method used is known to enhance the validity of qualitative research and that the codes were developed after a thorough review of the literature, I am

confident that a similar research study would generate codes and themes similar to those created in this study.

Role of the Researcher and Researcher Bias

Arguably the most important component to consider when designing a qualitative research study is the biases of the researcher. In qualitative research, the researcher is the research instrument. Merriam (2009, p. 17) listed the following competencies for a qualitative researcher:

- A questioning stance with regard to one's work and life context
- High tolerance for ambiguity
- Being a careful observer
- Asking good questions
- Thinking inductively
- Comfort with writing

Lincoln and Guba (1985) and Patton (1990) argued that critics of qualitative research contend that the research is more dependable than it is reliable, and that difference stems from the researcher serving as both the data collection instrument and the analysis instrument. Due to these concerns, it is important that the research follow a plan of inquiry to stem biases and produce reliable results (Sanjari, Bahramnezhad, Khoshnava Fomani, Shoghi, & Ali Cheraghi, 2014; Merriam, 2009). As a long-time employee within the supra-district that the Seven Fountains district lies within, it was important that I examine my own biases because they could, and in fact without examination and identification of those biases and perceptions probably would, impact the results of the

study. This process of reflection and identification is known as reflexivity (Darawsheh, 2014).

For myself, I engaged in consistent, deliberate, and meaningful reflection throughout the study in order to identify my biases. For example, I needed to acknowledge my own thoughts on teacher evaluation as both a former teacher and a current administrator. I needed to reflect that the teacher evaluation process for me as a teacher left me wanting, and that as an administrator was daunting. My view of evaluation, and specifically of teacher evaluation, was changed by the review of the literature that I underwent as a part of this study. I grew in both my understanding of the usefulness of appropriate evaluation and also the close connection that the fields of evaluation and education have. The linkage between these two fields gave me a new appreciation for the mutual goals of the fields and the nature in which they evolved together and overlap.

In reflecting on the interviews, I needed to ensure that I did everything possible to remove my perceptions or biases from the process of conducting the interviews and analyzing the data. To do this, it did help to have the expert panel review the questions to ensure that none of the questions were leading questions. I practiced delivering the questions in a neutral tone of voice and in the same manner repeatedly. I worked with the expert panel to anticipate some potential follow-up questions in these semi-structured interviews so that I would practice avoiding leading follow-up questions. I needed to acknowledge that although I do not work in the Seven Fountains district, I had worked in an affiliated district and therefore had met some of the participants in collegial settings prior to the commencement of the study. I needed to be aware that because that made me

an “insider,” the participants may not fully answer every question assuming my prior knowledge about the context, policies, and practices of teacher evaluation in Seven Fountains. Despite knowing some of the participants, and all participants knowing I come from an affiliated district, I was not a supervisor. I think that my familiarity with the teacher evaluation context, policies, and practices of Seven Fountains, coupled with my familiarity as an “insider,” would enhance the trust that the participants would have in me as the researcher. To further solidify the participants’ trust, I obtained signed informed consent forms that guaranteed the participants’ anonymity and their understanding of what information would be gathered and reported for the study.

No researcher can guarantee absolute objectivity in a study; however, the methods employed in this study, along with the reflexivity described above and in the Background section of this chapter, are important in reaching near complete objectivity. My purpose in conducting this study was to discover principals’ perceptions of the use of the Danielson *Framework* in improving their instructional leadership. After reviewing the literature on qualitative methods, I was able to design a study that would effectively fulfill this purpose.

Summary

Chapter III provides a comprehensive explanation of the methods used in this study, why those methods were chosen, and how they were enacted in this study. It included the reasoning and means for the structure of this study. The next chapter presents the results of this study’s methodology.

CHAPTER IV

FINDINGS

Introduction

Chapter IV reports the findings and an in-depth analysis of the findings. The Introduction provides an overview of the purpose of the study and the four research questions that guided this study. The remainder of the chapter provides the in-depth analysis of the findings by research question.

The purpose of this study was to examine principals' perceptions of their ability to effectively use Danielson's *The Framework for Teaching* to provide formative feedback to teachers, make summative decisions on tenure and termination, and improve their instructional leadership of the school.

Changes and added complexities to teacher evaluation represent a significant shift in the principal's responsibilities. Although this shift began in the late 1980s, the recent decade has seen a dramatic acceleration in combination with a considerable array of non-instructional, but still vital, principal responsibilities (Fullan, 2014). This study sought to understand how the principals used *The Framework* to become more effective instructional leaders of their schools in light of the recent implementation of Danielson's *Framework*.

Twelve principals from the Seven Fountains School District participated in a semi-structured interview composed of questions specifically designed to address the four research questions that guided this study:

1. How do principals perceive their preparation to use the Danielson *Framework*?

2. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is effective in providing formative feedback to teachers?
3. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is beneficial to the tenure and termination process?
4. To what extent, if any, do principals perceive that the use of the Danielson *Framework* has improved their instructional leadership?

These research questions served as lenses in the development of significant themes related to principals' experiences with the use of the Danielson *Framework*.

Themes Related to Research Question 1

Research Question 1: How do principals perceive their preparation to use the Danielson *Framework*?

Introduction to the Findings Related to Research Question 1

Before discovering the principals' uses or perceived effectiveness of the Danielson *Framework*, it was important to discover the principals' perception of the preparation to accurately utilize the Danielson *Framework* observational instrument. Three themes emerged from the conversations with principal participants:

1. Differences from past experiences
2. Self-directed preparation and training
3. Variability in ratings and relevance

In discussing the difference from past experiences, the principal responses indicate that principals consider the Danielson *Framework* to be different from their experiences as a teacher. Principals' responses indicate that their preparation and training were largely self-directed (reaching out to peer principals, individual research, and seeking training

outside of the district). Principals' responses indicate a lack of alignment and common understanding of the *Framework's* rubric rating levels (Highly Effective, Effective, Developing, and Ineffective), and concerns that the *Framework* is not relevant to all subjects and grade levels.

Differences from Past Experiences

Ten of the 12 principals indicated that the Danielson approach to evaluation was generally different than their experience as a teacher. Six of 12 principals indicated that the Danielson model was entirely different. In discussing his experience with evaluation when he was a teacher, Principal A stated, "Mine was pathetic. It had questions that said, 'Is the room well ventilated?' 'Is it neat and attractive?' 'Are the blinds open to let sunlight in?'" Principal F agreed that his experience had also been this checklist-styled approach to evaluation, citing differences between his experience with evaluation and teacher evaluation under Danielson by stating the following:

"Night and day. It basically was a checklist very, very—it wasn't structured this way . . . It was just more of someone sitting in the back of the room giving all "G's" ["G" = Good] at that time, and no discussion out there."

Other principals agreed with strong, clear answers about the differences between past evaluations and evaluation with the Danielson *Framework*; for example, Principal H stated, "Radically changed . . . Totally different . . . So different." Principal D stated, "Totally different," and Participant G stated, "Completely different."

Seven of the 12 principals indicated that the use of the Danielson *Framework* brought more specificity to the evaluation process. Principal E concurred with those statements when describing the evaluation model in use when she was a teacher as

“More generalized, and this [the Danielson *Framework*] is much more specific although all those things . . . may have been in the evaluators’ minds; they were not necessarily written down, and having to give those four category responses, it is like “Yes, she did, you know. She didn’t do that . . . Not really judging me on how I did it. So I think what we have [now] is a little bit more rigorous in that regard.”

Self-directed Preparation and Training

Five of 12 principals indicated they felt they had the necessary training and preparation to use the Danielson *Framework*. The remaining seven principals, felt they had low levels of preparation and training. Principal D summarized many of the seven principals by stating, “I find it more difficult to use, and maybe it’s because I don’t feel I was fully trained in it. And it’s just something, you know, very, very new to me.”

Principal A elaborated on the training he received, stating the following:

“I think I’d like to have had a little bit more. We had a pseudo training that lasted four hours. If I remember right, it was about four hours . . . where they took us through what each one [domains and ratings] stood for. I think I would have appreciated it more if we had a video of somebody teaching the lesson, then critiqued it, and then went over our review. I think that would have been much more beneficial. I’ve learned from my own experience through when I’ve gone back and I’ve actually changed what my initial rating was after I really thought about it or saw it in comparison to something else.”

Four of 12 principals discussed how the Danielson *Framework* did not apply, or

not easily apply, to all teachers of every subject and at every grade level. Principal H concurred with the seven principals, stating “I feel that I haven't had enough training and I actually did go for training for it [the Danielson *Framework*] at Iona College; and I don't feel that I have all the domains, all the pieces A, B, C, D, E, F, in place . . . To effectively observe a teacher.”

Eight of 12 principals indicated that they spent time teaching themselves or working with principal colleagues to improve their ability to use the Danielson *Framework*. In essence, these eight principals are indicating that they are informally training with self- and peer-direction. Principal F encapsulated it best for the eight principals when stating, “You just have to do your own homework in your field to promote yourself and. . . rely on others.” Principal B concurred that the training is self-directed, stating, “I have only gone through my own readings and understandings.” Principal G explained how colleagues can help by stating the following:

“I have a lot of conversations actually with a colleague principal. We will talk . . . I'll show him something that I've written, and I'll ask him to help me to clarify my thoughts. He has never actually seen that lesson, but I think there would be use in having that done . . . I wish that we could be co-evaluators . . . he could, because he is strong in math, he could come here and evaluate. And where he is weak in ELA, I could go there and help him there.”

Principal G's desire to be side-by-side for support in evaluations was substantiated by Principal F's self-directed experience:

“I personally spent a lot of time researching it beyond going online,

I sat with principals who work in both public and private schools and had a conversation where they brought me into their world. Some of them actually walked me through their building, and their classrooms, and told me how they did it and why. And that was the best experience I had.”

Variability in Ratings and Relevance

Only one of 12 principals indicated that their understanding of the ratings was aligned with their principal colleagues, and five of 12 principals indicated that ratings were not aligned among colleagues. The remaining six principals felt that there was not a strong alignment. Participant H explained as follows:

“I don't think that we had a workshop or a presentation where we have addressed different domains and all the subtopics of those domains as a district. So, therefore, I'm sure one principal will be evaluating, or an assistant principal will be evaluating, a teacher very differently than another one . . . It all has to do with the mindset and the amount of knowledge you have of the Danielson because it's very different than what we had previously.”

Participant A verified the actual differences in the ratings as follows:

“I think I have a more structured belief in Developing. I don't believe a Developing should always be Developing. And I feel that if it doesn't go from a Developing to an E [effective] and it remains that Developing. By the end, it really is Ineffective when it comes to the total picture. Most of my cohort principals don't feel that way. They feel that you can maintain Developing. I don't.”

Principal E explained the reasoning for the lack of alignment with ratings, connecting back to training and experience. She stated the following:

“I think that perhaps principals would have had more experience and might look at things differently than the principals that don't have as much experience. And principals who excel in that subject, whatever subject you're observing, would also have different insights. And principals who taught in the classroom, maybe that grade and maybe even lesson. It varies by who the principal is. So I would say that it's not going to be the same.”

The lack of alignment in ratings can even manifest within the same school.

Principal K succeeded a prior principal and could not trust the ratings from the previous principal “because the previous principal may have given them glowing reviews, and I came in with a very different perspective, and so they didn't link up.” In explaining variability within her own building, Principal I stated, “That teacher is at a lower grade so the expectations are not the same as what I'm expecting you [an upper grade teacher] to do in your classroom and what I expect them to do in their classroom; there is a difference.”

Principals, B, E, and G all used the word “subjective” when discussing ratings alignment. Principal J encapsulated the feelings of subjectivity when stating, “I know there are principals who give HE's [Highly Effective] to everybody down the line. So I think that any tool is as good as the person who's using it.” Principal D furthered Principal J's statement about the [perspective] of the principal using the tool, saying that “As in anything else, our expectations can be different. Some will be similar to us and

some will be dissimilar.” In summarizing the subjectivity, Principal G stated, “So I would say, yeah, it's subjective and what I feel could be . . . that somebody could Develop, and because I always give people the benefit of the doubt, [some] people may just [go] immediately to Ineffective.”

Zero of 12 principals stated that they had participated in calibration observations with their fellow administrators. Participant H added, “I don't think I have at all. I'm in the district for 34 years.”

In addition to the variability in ratings, four of 12 principals also discussed the variability in the relevance of the Danielson *Framework* in evaluating all grade levels and subjects.

With regard to this, Principal G stated the following:

“I have mixed feelings on it because unfortunately the evaluation is being used for Pre-K teachers through Grade 8 teachers and in various subjects. When I am using the evaluation for [a] Pre-K teacher, many of the fields are not valid. Even in some Level 1, Grade 1, fields. So it's difficult for me to fairly assess them, and I end up skipping a lot of things on the form. Also, I feel the form doesn't lend very well to math.”

Principal I furthers this point regarding subjects, stating “For my special teachers, it really is very difficult to use . . . You know there's certain things in it that don't apply to them . . . So, I find it very difficult. I think the evaluation should be different.” Principal L stated that with regard to different grade levels, “It doesn't seem to suit the needs of early childhood instruction the way it is set up now.” Principal I succinctly stated the

concern among some principals when stating, “I think a lot of the items on it [the Danielson *Framework*] are not always relevant to the lesson.”

Summary of Findings Related to Research Question 1

Discussions with principals regarding their preparedness indicate an overall lack of preparedness to effectively utilize the Danielson *Framework*. Principals’ indications that the *Framework* was significantly different than their experiences as a teacher, coupled with a feeling of greater formal training and no calibration experiences, logically leads to the principals’ beliefs that ratings are not aligned and the *Framework* is not universal.

Themes Related to Research Question 2

Research Question 2: To what extent, if any, do principals perceive that the use of the Danielson *Framework* is effective in providing formative feedback to teachers?

Introduction to the Findings Related to Research Question 2

Improving instruction is an iterative and formative process, and any teacher observation model needs to be useful in providing formative feedback. Three themes emerged from the conversations with principals regarding the formative uses of the *Framework*:

1. Useful tool in providing formative feedback and coaching
2. Focused conversations about teacher practices
3. Relationships and personalities

Principals indicated that the *Framework* was useful in providing feedback because the *Framework* provides principals the ability to identify instructional needs and then provides the next steps for coaching. Principals felt their conversations were now focused

on instruction and teacher practices because the *Framework* provides that focus on the teacher practices. In working with teachers, principals discussed how relationships were generally improved between principals and teachers but that impact on a teacher's instructional practices were largely dependent on the teacher's personality.

Useful Tool in Providing Formative Feedback and Coaching

"I think I've focused a lot more on growth," as stated by Principal J, well summarizes the 9 of 12 principals who indicated that they find the Danielson *Framework* to be useful in providing formative feedback to teachers. Principal L spoke of the usefulness of the *Framework*, stating, "I have *actual feedback* that I could show them as to what I mean when I say I want to improve instruction" (emphasis added by the researcher). The "actual feedback" is an indication of the usefulness of the *Framework* in comparison to the previous observation model used in the district. Principals B and F concurred that the *Framework* was helpful in providing targeted feedback. Principal B stated, "It allows me to look to be very targeted in terms of specific areas . . . whether it's instruction, planning, or environment . . . I'm able to then say and use this as a data point with the teachers. 'It's there' . . . 'it's not there,' etc.," and Principal F stated, "Rather than saying, 'I didn't like your room; change it,' I gave her reasons why. And this helped me sound more professional as to why. I was linking it to something."

Principal K spoke of the *Framework* as beneficial in seeing where to make improvements, stating, "It really does help me in seeing what I have and what I need to work on with the teachers." Principal K reflected how now having the ability to identify areas of need and "being able to have conversations about them and develop . . . plans and procedures to improve."

Principal F likened this newer role in providing formative feedback to an athletic coach, stating, “We are really diving in deep, as though I was a pitching coach. Pitch one, Steve, I’m going to talk to you about your form.” Eight of 12 principals indicated that they used the Danielson *Framework* for moving beyond simply formative feedback and into instructional coaching based on that feedback. Principal E explained how using the *Framework* allows for incremental coaching, stating, “So we can focus on one or two in a conversation and then say, ‘Now how are we going to build it into your lesson?’ So there’s coaching there happening—let’s build the lesson together so that we can see how you can do this, and hold your hand while, you know, you’re doing it.” Principal K also stated the *Framework* is helpful with individual coaching. “I’ve used it more just for individual one-on-one kind of professional development of—you know, more of the coaching aspect.”

Seven of the eight principals who discussed using the *Framework* to improve their instructional coaching discussed working with individual teachers. For example, Principal G stated, “But when I’m coaching them . . . when I’m giving them that advice, it’s kind of a brainstorming session together. Let’s talk about what happened because what works for me may not work for them.” Principals A and F discussed modelling as coaching. Principal A stated, “It allows me to actually go in and model what I’m looking for and they are looking forward to that.” Principal F said, “And I’m working with them a lot in the beginning, particularly the ELA teacher, to actually show how I would do it. Kind of do a lesson or two for them and then they take it from there.”

One principal, Principal I, did not speak of the principal-teacher approach to coaching but did still emphasize the use of the *Framework* in informing instructional

coaching, stating, “I hire people to come in and do coaching. I look at their observations and my walkthroughs and what I see on a daily basis. And I try to fill in those gaps with people who have more expertise than I do in particular areas.”

Principal G was mixed about the use of the *Framework* for formative purposes, stating, “I have four meaningful conversations with the model present and in front of me, but the rest of the conversations are not with the model. I may refer to the model. I may refer to aspects of it or you know an area that we need to focus on, but I don't use it on . . . it's not something I would use on a regular basis, no . . . So my post observation conversations are beneficial using the model, but on the day-to-days I don't do that.”

Focused Conversations about Teacher Practices

Seven out of 12 principals indicated that using the Danielson *Framework* provided an instructional focus either for them, for the teachers, or for both. Principal G stated, “I understand and appreciate Danielson now, having been an administrator because it helps me have some of the tough conversations with teachers.”

Reflecting on the teacher evaluation model prior to using the *Framework*, Principal A continued, “It [previous model] had nothing to do with instruction. No, this one [Danielson *Framework*] absolutely forces the observer to look at plan and practice.” Principal E also used the word *forces* when discussing the focus on instruction, stating, “Makes you think about things that you might not have thought about; it forces you to look at things that you might have glanced over.” Principals B and D both discussed how the *Framework* provides a focus for observing instruction. Principal B stated, “Helped me to define specific areas that I want to look for.” Principal D stated, “I think it made all of us take a second look because of the detail.”

The instructional focus provided in using the *Framework* prepares principals for instructional conversations with teachers. Principal A framed the difference of an instructional conversation when using the *Framework*, stating, “I think a lot of times we have a habit of just talking to talk and feedback gets lost. This keeps us on track as ‘This is what I want to talk about because this is what you're held accountable to.’” Seven out of 12 of the principals indicated that using the *Framework* provided a starting point to conversations with teachers. Principal F stated, “This is my starting point. This is how I want the classroom to be. What is your vision? Where do you feel you've struggled? Where are your successes and why? And what did you do for those children who you felt were struggling?” Principal K stated, “Even in terms of that it gives me . . . when I'm going into the classroom . . . it gives me kind of a starting point in my head like, okay, this is what I'm looking for—am I seeing it? Or, you know, am I seeing it and seeing it done correctly? Am I seeing it done well? Or am I not seeing at all? So, kind of gives me that baseline.” “It's a conversation opener,” stated Principal H, which was similar to Principal A, who stated, “I think it's the start of many of our conversations” and to Principal K, who stated, “So, like I said, it just gives us kind of a starting point of what we want to talk about.”

Relationships and Personalities

Eleven principals of 12 discussed relationships with their teachers since using the Danielson *Framework*:

- Two of 12 principals indicated that there was no change in relationships and that they were positive overall.

- Five of 12 principals indicated that some relationships with teachers were better, and some were worse.
- Four of 12 principals indicated that relationships had improved.

Principal A stated, “Half of them will make their own appointments to say ‘Gee, I’d like to follow up with . . . or what do you think?’ So it opened a really nice door to those. Those that see it as an attack never bring it up or talk about it.” Principal F echoed the experience of Principal A, stating the following:

“Not everyone here . . . I have a couple that ‘get it’ but feel a different methodology would work. I have two that really give resistance. They give resistance when we have a general conversation about a walk through, when we have a formalized evaluation talk, even in passing.”

Principal E stated, “For some teachers, the ones who get the Highly Effective, we have a great relationship. But for those who have a D [Developing] or, you know, an I [Ineffective], it makes it very difficult.”

Principals H and K reflected how initial instances using the Danielson *Framework* can be challenging. Principal H stated, “So it did, early on, impact in a very difficult, negative way.” Principal K stated, “And so it created a little bit of tension that this teacher who had been doing it for so long . . . here’s a new principal who came in and gave her a Developing.”

Six of the 12 principals talked about how the impact of the *Framework* on teacher practice is influenced by the personality of individual teachers. Principal H said, “I think it depends on the teacher. It’s different.” Principal A said, “Some of them see it as a collegial conversation; others see it as an attack.” Principal E stated, “So, you know, try

to be as gentle as you can; but it doesn't always get received in the way that you want it to, depending on personality and depending on the previous relationship with the person.”

Principal E continued, saying, “But that depends on their personality because not everybody takes criticism well. Or feedback well.”

Principal F discussed personality in terms of a teacher’s openness to change:

“On the flipside, I have a teacher that doesn't have a buy-in. It becomes a challenging conversation because they ask ‘why?’ a lot. ‘Why do I have to do that, this has worked in the past?’ I bring them a lot of research that is tied to Danielson, but it doesn't go anywhere because there's no buy in.”

Principal F continued, stating, “I feel that they have not grasped it; their rationale is that they’ve been teaching 20, 25, 30 years and some things relate, and some don’t. They feel that sometimes how it used to be is better.”

Summary of Findings Related to Research Question 2

Despite the impact of a challenging teacher personality, principals are using the *Framework* to improve instruction. In using the *Framework*, principals are able to improve relationships with teachers, have targeted and focused instructional conversations, and work as instructional coaches to impact instructional practices.

Themes Related to Research Question 3

Research Question 3: To what extent, if any, do principals perceive that the use of the Danielson *Framework* is beneficial to the tenure and termination process?

Introduction to the Findings Related to Research Question 3

A central evaluative decision for principals is whether to offer tenure to a teacher. Offering tenure essentially means that job performance so far was sufficient enough to guarantee that job, nearly, for life. Two themes emerged from the conversations with the principals:

1. Clear expectations for job performance
2. Evidence to determine teacher tenure

The *Framework*, according to the principals in this study, provides clear expectations for instructional practices and, therefore, a teacher's job performance. Additionally, the principals indicated that the *Framework* provides sufficient evidence to both make an informed decision about tenure, and to support any challenges to that decision.

Clear Expectations for Job Performance

Nine of 12 principals indicated that the Danielson *Framework* provides clear expectations for instructional performance. Principal A stated, "I think there are real conversations now. And they're grounded . . . in fact, they're grounded with a lens that both of us know, and it's not something that I've decided to grab from the air today because I have nothing better to do. Principal L continued this common understanding of expectations referred to by Principal A, saying the following:

"From the outset when we have our September meeting, they do get a copy of the form; and it's also in the union contract book. And those are the expectations set out when they plan their lessons. And what I expect to see when I walk into the classroom, given frequent observations, [is] evidence of what you're teaching, how your instructional practices [are] going."

Principal G stated, “I know what I'm looking for overall in their regular observations and then their annual performance appraisals.” Principal C supported that idea, stating “I feel that because the *Framework* is so complete in terms of the way it is written, it's very clear when you're having the discussions about what you would like to see, if there's anything different that you'd like to see, or what you have seen that you appreciate, it makes it easier to talk about.” Principal C continued, “In the same way that, you know, that we establish rubrics for our students to hopefully help them clearly understand what the expectations are, I think this basically works the same way.” Principal C continued and encapsulated the clarity in the summative evaluation of teachers since changing to the *Framework*:

“I just feel like sometimes in the past, before we went to the *Framework*, you would get pushback on certain things; and I think some of that was vocabulary issues that have been mentioned before. And some of that I just think that, you know, people would want to defend what they were doing; and because we didn't have this *Framework* in place . . . made that a little bit easier. So, there could be a lot of pushback. Like I said, [with] this, I find that it's more productive and that it's very clear what I'm looking for, and expectations are so much more clear. And then at the same time, for them to give me the feedback, it's also very clear that what they're thinking about when I'm expressing to them what I've seen . . . what my evaluations are.”

Evidence To Determine Teacher Tenure

Principal A stated that with the use of the Danielson *Framework*, tenure expectations . . . “They’ve risen.” Principal F explained the heightened expectations by

saying, “We're looking under an academic microscope now, not just procedural . . . you got three years.” Principal K stated, “As an evaluator it [the *Framework*] helps me to see their effectiveness.”

Twelve of 12 principals discussed how the *Framework* has provided evidence that can help make summative job performance decisions such as recommendations for tenure. Principal L stated the following:

“It's changed a lot because now the categories are very specific. You either do it or you don't. You're proficient or you're not. And if you're not proficient, you're working towards it; so then it gives you a basis for conversation and to say you're not ready you're not there yet. Maybe you will be next year but this year you're not ready.”

Principal L continued this point about the need for teachers to demonstrate improvement, saying, “If those changes aren't implemented, then I can go in as an evaluator and say, ‘You're not proficient in this area.’” Principal I agreed with Principal L, saying she was also looking for improvement:

“I think that it's helped me a lot; I think the old evaluation missed a lot of points, and the Danielson has a lot that I personally looked for in my teachers. And [it] has been able to allow my teachers to grow, and helped me to know if they should be tenured, or not, or extended another year.”

Principal H added that the process of measuring and looking for improvement was a more data-driven process to evaluation:

“I think now you do have those specific areas where, you know, we have to get multiple observations where in the past you didn't.

So, you know, if you give your first observation in 45 days, you can give feedback. And then . . . giving another observation within the next month, you have more documentation on that teacher as to whether they are improving, becoming more effective or not. So, yes, it's data-driven.”

Principal F stated, “Tenure has changed now. In the past, it was almost automatic; now it has become more of a deeper process.” Principal J added to the changing nature of the tenure process:

“I think it's [the *Framework*] much more beneficial in the sense that when I came in as principal and I looked at everybody's observations before we moved to Danielson, every single person in the school was a good teacher. They all had ‘Good, Good, Good, Good, Good,’ and people had tenure that should not have had tenure. And here you can look at somebody and without dinging them, I mean they might have some D's [Developings]. They might have, you know, they're Developing . . . they're Effective. But you can say, ‘This isn't where you should be at this point,’ and I think it allows for more.”

Principals E and I emphasized the importance of evidence. Principal E stated the following:

“It has provided documentation that may, or may not, have existed before as to why this person . . . I would recommend this person, or I wouldn't recommend this person. And I think that's important. Again, it's not easy, if, especially, you're not recommending

the person. But it is documentation that is needed so that you can show why you asked this person repeatedly, you know, ‘These areas need to be worked on,’ and it hasn't gotten done.”

Principal I concurred, stating, “It's back-up for me.”

Summary of Findings Related to Research Question 3

Principals’ statements indicate that the tenure process and summative evaluation of teachers’ performance has changed since using the *Framework*. The tenure process used to be more procedural; for example, after a certain number of years teaching, a teacher was granted tenure. The use of the *Framework* has changed the process of tenure to be about the quality of a teacher’s instructional practices. Principals indicated that the *Framework* provides both the clear expectations for performance and evidence needed to make the determination for tenure.

Themes Related to Research Question 4

Research Question 4: To what extent, if any, do principals perceive that the use of the Danielson *Framework* has improved their instructional leadership?

Introduction to the Findings Related to Research Question 4

The principal’s role in instruction is of paramount importance. The *Framework* provides principals with a tool to help monitor and improve teachers’ instructional practices. Three themes emerged from the discussions with principals:

1. Improving instructional leadership practices
2. Setting high expectations
3. Principals as leading learning

In using the *Framework*, the principals in this study indicated that they are better instructional leaders. Principals are able to use the *Framework* to set high expectations for instructional practices. Fullan (2014) indicates that leading learning consists of three components (capacity building, collaborative cultures, and learning communities) and the principals' statements were mixed about how the *Framework* has contributed to their efforts in these three areas.

Improving Instructional Leadership Practices

Six of the 12 principals strongly indicated that they are better able to be an instructional leader since using the *Framework*. Principal A stated the following:

“I'm in the classroom a lot more. I think when you really look at the breadth and depth of Danielson in a 45-minute, 60-minute lesson, it's really not enough to give that that clear of a picture to really hit all of that. So, while, yes, I have to use that in that 60 minutes or 45 minutes, it's really framed all of my walkthroughs, my thought processes. Why am I doing what I'm doing? Why are they doing what they're doing? How is it helping the kids? Is it helping the kids? I'm not looking for the superficial—what's on the walls—any more. Yes, anchor charts are great; but if you have the best anchor charts in the world and the conversation going back and forth is nothing, the aesthetics and all which should be in a classroom are not going to do anything. We call them, again, targeted walkthroughs, but it's really targeted looking at instruction.”

Principal K added to Principal A's statements on being able to go deeper into evaluating instruction:

“it helps me in knowing where teachers’ strengths and weaknesses are and identifying them and then being able to have conversations about them and develop plans and procedures to improve. Or sometimes, you know, with strengths, share them with colleagues. You know, on second grade math if you’re on a break, pop in there; just watch how it operates. So, it really does help me in seeing what I have and what I need to work on with the teachers.”

Principal H also talked about being about to lead instruction more effectively due to deeper instructional leadership skills:

“Because looking at what an exceptional teacher would do in a classroom based on each of the topics in each domain, it has helped me understand teachers who need a little bit more support. I mean if you look at the breakdowns, it really does give you an example of what this type of teacher does. So, I think it’s informed to me.”

Principal F stated, “This [the *Framework*] provided better language for that, this provided a framework to put it in, it gave me containers. Virtual containers and virtual tools to talk about it with them. It wasn’t just sporadic; it’s better to have a focus.” Principal H stated the following:

“I think it’s due to the *Framework* because if you understand the *Framework*, it really does address everything we should be doing for our teachers and for our students. It’s extremely comprehensive. I think it has helped me address, especially, depth of knowledge in the classroom with the teachers. I think I have more confidence now as a coach.”

Principal B furthered this point, stating, “It allows me to be very targeted in terms of specific areas within . . . whether it's instruction, planning, or environment, I'm able to then say and use this as a data point with the teachers. ‘It's there,’ ‘it's not there,’ etc.’” Principals C, H, and E furthered Principal B’s statement about the ability to target by discussing the specificity of the *Framework* and its usefulness. Principal C stated, “But I think it does a much better job of breaking them down into their discrete pieces so you can analyze each one individually.” Principal H added, “In the domain you can be really good in certain areas, and areas where you need a little . . . I think everybody does it; nobody's perfect really.” Principal E continued, “So we can focus on one, or two maybe, in a conversation and then say, ‘Now how are we going to build it into your lesson so there's coaching there happening? Let's build the lesson together so that we can see how you can do this.’”

Principals J, F, and I discussed the impact of the *Framework* on macro-instructional areas. Principal J stated the following:

“I've focused a lot more on looking at trends within the faculty, which . . . I like the Danielson *Framework* for that because I can look at . . . So, I am a little bit nerdy when it comes to statistics and things; and after I've done my walkthroughs and after I've done my formals, I kind of take them and sort them to see where . . . [they] were struggling. And that really speaks to a good conversation in our faculty room . . . noticed that, you know, nobody has [particular] stuff on the wall or, you know, stuff like that. Or I notice that we do really well with this but we don't . . . and then it becomes a conversation of how we can fix that . . . I like the way it's broken up.

Because I feel like often teachers excel in one domain and we'll struggle in another . . . which I like to see. Again, I love patterns.”

Principal F concurred, stating, “It's given that professional language based on that shared vision.” Principal I stated that the *Framework* “determines what programs are going to change in the school.” Principal I continued as follows:

“I see through Danielson where the teachers are lacking and I do compare it to test scores (ITBS, state scores, Fountas and Pinnell) so it's a data point. Then I compare all of those to determine where my coaching might be going. Who I'm going to bring in.”

Setting High Expectations

Seven of 12 principals indicated that their expectations for teachers' instructional performance has increased since using the Danielson *Framework* for teacher evaluation. Principal E stated, “Well, it's helped me to be more rigorous.” Principal G stated, “I definitely think I'm a lot harder than I used to be.” Principal G continued, stating the following:

“And I think that when I go in there, like, ‘You're not doing this, you're not doing that.’ I'm okay, you know, ‘You're not doing it, you're not doing a lot of things. I would be okay you're having an off day; but you know we all have those, and these kids still need you. So how are we going to fix it?’ So, I guess, in that way, I'm a little harder.”

The high expectations of the *Framework* are seen by some principals as a shared goal. Principal A stated the follows:

“They [the *Framework* standards] certainly set the bar high for the

teacher to achieve. But it certainly allows the principal an established set of criteria that you know what you're being looked at for. It's your responsibility to work towards it. And it's a shared responsibility I think, at that point to mend what's not working.”

Principal J stated, “I think it's really helped me look at it as a growth model and making even the best teacher better. And then, on the other hand, having a strong framework for a struggling teacher.”

Principals as Leading Learning

Principals provided a wide variety of responses related to the use of the *Framework* in providing professional development:

- Three principals indicated they did not use the *Framework* at all in planning professional development
- Three principals indicated they did use the *Framework* (rather extensively) when planning professional development,
- The remainder of principals indicated they somewhat used the *Framework* in planning professional development.

Principals C, F, and K all indicated they did not use the *Framework* in designing professional development, stating the following: “You know we really haven't” (Principal C), “I don't personally design professional development” (Principal F), and “I have to say I haven't used it to implement whole staff development” (Principal K).

Principals I, E, J, and H all indicated purposeful use of the *Framework* when designing professional development. Principal I stated, “I look at their observations and my walkthroughs and what I see on a daily basis. And I try to fill in those gaps with

people who have more expertise than I do in particular areas.” Principal E stated the following:

“So there are five people that have a D in one area. That's a good indication to me that people are struggling with that concept. So what can we do in terms of getting professional development on a half day when we have teacher meeting/teacher PD? Somebody can come in and talk to them . . . We could watch a webinar and discuss it.”

Principal J added, stating the following:

“And even that something as silly as I did a bunch of a walkthroughs this last week . . . I always do an article of some type in my weekly newsletter to the faculty and the one I've picked is something that I noticed people weren't doing. So, they may not notice that I'm doing that, but that's where I will, in a faculty meeting . . . I'll say you know we really need to work on questioning.”

Principal H continued discussing the use of faculty meetings, stating, “Every faculty meeting [I] try to do some sort of P/D. All right, so right now we're taking data from MAP, from the state scores, and we're looking at how they correlate and how does it correlate with your classroom. So, triangulate all the information.”

Seven of 12 principals discussed their efforts to build capacity, collaborative cultures, and/or learning communities. These three efforts are identified as components in Fullan's (2014) leading learning theory. As an example of building capacity, Principal G discussed the actions of her instructional coach, stating, “So she [the coach] will, you know, do a little P/D on how to do that and what it looks like . . . and then one-on-one

with a teacher where she's modeling and guiding them.” As an example of collaborative cultures, Principal L stated, “We try to work collaboratively so that as the children go through each grade, they're not relearning certain practices. They're going from one teacher who is teaching writing a certain way to the next teacher who will enhance writing based on what was learned the year before.” Principal F also discussed collaborative culture building stating the following:

“My biggest thing is to foster collaboration; we don't have multiple Grade 1's and Grade 2's here, but we have departmental settings and we have teachers now where K will talk to 1, and 1 will talk to 2, and they do share activities together. This has grown over the last couple of years in this school, and it's because we now share a vision that's more structured.”

Principal F also provided an example of learning communities, stating, “But this is typically what we started in September or the June before. It's always at the September meeting and then at the June meeting where we have it all written out—I can even show you—and then we go over it. What did we learn? What didn't we learn?”

Summary of Findings Related to Research Question 4

The principals' statements indicate they have strengthened their ability to be instructional leaders in their schools. In using the *Framework*, principals are able to set high expectations for teacher performance, identify when those expectations are not being met, and work to improve those areas of need. Principals' responses indicated that the *Framework's* specificity of language is helpful in improving their instructional leadership.

Summary

The purpose of this study was to examine principals' perceptions of their ability to effectively use Danielson's *The Framework for Teaching* to provide formative feedback to teachers, make summative decisions on tenure and termination, and improve their instructional leadership of the school. The themes that were uncovered during the analysis of these data, and previously presented in this chapter, provide data that directly address the purpose of the study.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Introduction

This chapter presents an overview of this study. This chapter begins with a summary of the purpose of the study and a restatement of the research questions. The chapter continues with a discussion of the findings of the study within the context of the existing literature and the lenses of the theoretical frameworks. The chapter concludes with considerations and recommendations for practice, policy, and future research.

The purpose of this study was to examine principals' perceptions of their ability to effectively use Charlotte Danielson's *The Framework for Teaching* to provide formative feedback to teachers, make summative decisions on tenure and termination, and improve their instructional leadership of the school.

Changes and added complexities to teacher evaluation represent a significant shift in the principal's responsibilities. Although this shift began in the late 1980s, the recent decade has seen a dramatic acceleration in combination with a considerable array of non-instructional, but still vital, principal responsibilities (Fullan, 2014). This study sought to understand how principals used *The Framework* to become more effective instructional leaders of their schools in light of the recent implementation of Danielson's *Framework*.

Teachers and principals represent the most significant, controllable variable on student achievement (Baker et al., 2010; Darling-Hammond, 2000; Tucker & Stronge, 2005; Taylor & Tyler, 2012). This study therefore focused on the critical components of student achievement and how Danielson's *Framework* contributes to those components of the evaluation process from the perspective of principals. Principals' ability to measure

and help improve teaching is critical. As Stufflebeam and Shinkfield (2007) stated, “Society has a critical need not only for competent evaluators but for evaluation-orientated decision makers as well” (p. 5).

The ability to continually support teachers throughout their careers and to identify teachers who could serve as coaches or other instructional leaders to their peers would be invaluable to schools and to the field of education (Darling-Hammond & Ducommun, 2010). Ample research supports the use of classroom observations to support instructional improvement (Gallagher, 2004; Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004). Since classroom observations comprise a significant portion of *The Framework*, it is important to publish findings that may explain why early data from the new evaluation models, including *The Framework*, have not seen much change in teacher ratings (Keesler & Howe, 2012; Sawchuk, 2013; National Council on Teacher Quality, 2017). Due to this seemingly disjointed phenomenon, it is important for policy-makers and educational practitioners to understand why this is occurring.

This qualitative study consisted of a sample of 12 principals from the Seven Fountains School District. The research sample consisted of three male principals and nine female principals. The principal participants reflected a range of experience teaching, supervising instruction (not as a principal), and being a principal. While four principal participants taught for more than 15 years, the majority (7) taught for 10 years or less. The majority of principal participants have been instructional supervisors (8) and principals (9) for 10 years or less. All participants held a master’s degree. I conducted one-on-one semi-structured interviews with all 12 principals. The research was conducted in the spring and fall of 2018.

Research Questions

1. How do principals perceive their preparation to utilize the Danielson *Framework*?
2. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is effective in providing formative feedback to teachers?
3. To what extent, if any, do principals perceive that the use of the Danielson *Framework* is beneficial to the tenure and termination processes?
4. To what extent, if any, do principals perceive that the use of the Danielson *Framework* has improved their instructional leadership?

Summary and Discussion of Major Findings

Findings Related to Research Question 1

Research Question 1 one states the following: How do principals perceive their preparation to utilize the Danielson *Framework*? In the literature of the evaluation, The Personnel Evaluation Standards outline important requirements of the evaluation process: propriety standards, utility standards, feasibility standards, and accuracy standards (Joint Committee on Standards for Educational Evaluation, 1988). These standards relate directly to the data gathered in this research question. The utility standards require evaluators to issue results that are credible, informative, timely, and influential. The accuracy standards require that evaluations provide sound information about a person's qualifications and performance (with validly interpreted evidence). Understanding principals' perceptions of their readiness and preparation to effectively utilize the *Framework* observation tool would inform the extent to which principals would be able to meet utility and accuracy standards. As Stufflebeam and Shinkfield highlighted, there

is a “critical need not only for competent evaluators but for evaluation-oriented decision makers as well” (2007, p. 117).

The vast majority of principals considered the Danielson *Framework* to be considerably, even radically, different from their experience with observation and evaluation as a teacher. This is an important point because the newer the approach to evaluation, the greater the need for preparation and training to ensure that the tool is utilized correctly. However, despite the belief that the *Framework* was significantly different from their previous experiences, fewer than half of the principals indicated that they felt they had the necessary training and preparation to use the *Framework*, and all of the remaining principals felt they had low levels of preparation.

In discussing their preparation for using the *Framework*, principals discussed self-initiated learning activities and an organic reliance on fellow principals. Two-thirds of the principals discussed teaching themselves, or working with their colleague principals, to improve their use of the *Framework*. The cooperative nature of the eight principals, however, did not include calibration observations. No principals had participated in calibration observations as a component of their training (self-directed or otherwise). Due to the principals’ perceived lack of training, it is logical that 11 of 12 principals believed that their understanding of the rubric ratings were not aligned with their colleagues’ ratings. With regard to the evaluation literature, specifically The Personnel Evaluation Standards, 11 of 12 principals have indicated that their results are not meeting the utility standard for credibility and the accuracy standard for validly interpreted evidence.

Principals reflected that the *Framework* was much more specific and detailed than the tool used when they were teachers. Principals referred to the evaluation method used

when they were teaching as a checklist approach to evaluation and observation but agreed that the Danielson *Framework* was more comprehensive and detailed. This is consistent with the literature on teacher evaluation. The methods that the principals were describing were reflective of the clinical supervision approach to evaluation. Goldhammer's (1969) clinical supervision method of observation focused on teacher behaviors, but Danielson's *Framework* reflects a major shift based on an extensive review of the literature on teaching and psychology (Danielson & McGreal, 2000). Danielson (1996, 2007) states that The *Framework* claims to be comprehensive and generic to all disciplines; however, one third of the principals discussed the lack of relevance of the *Framework* to all grade levels and subjects. For example, the principals highlighted the lower grades (especially pre-kindergarten) and special subjects (for example, physical education) as areas without a direct, or at least easy, connection to the *Framework*.

Findings Related to Research Question 2

Research Question 2 states the following: To what extent, if any, do principals perceive that the use of the Danielson *Framework* is effective in providing formative feedback to teachers? The literature on student achievement proves the teacher to be the most critical factor in the academic success of students (Baker et al., 2010, Darling-Hammond, 2000, Tucker & Stronge, 2005, Taylor & Tyler, 2012). Having a quality teacher, therefore, provides the best pathway to strong student achievement, and knowing how to evaluate quality teaching becomes critical to student success. Principals are key to the success or failure because they are in a position to use evaluation as a formative process to be consistently improving teachers' instructional practices. While the literature proves teachers to be the most significant, controllable, factor in student achievement,

principals are the second most significant factor because of the influence they have over instructional quality (Leithwood, Louis, Anderson, & Wahlstrom, 2004).

Three quarters of the principal participants felt that using the Danielson *Framework* was useful in providing formative feedback to teachers. Participants discussed the ability to target specific aspects of a teacher's instructional practices as beneficial to the formative feedback process. Principals also discussed how the *Framework* provides them with examples of varying quality levels (principals seemed to, most especially, appreciate exemplary or "Highly Effective" practices). These feelings by the principals are logical, especially in light of their experience with the previous, less comprehensive, observational model they have used.

Eight of the principals discussed using the *Framework* as an instructional coaching tool. Principals discussed, extensively, the use of the *Framework* in working with individual teachers (such as in the instructional one-on-one coaching model). The ability of the *Framework* to support specific instructional practices, and at varying degrees of quality, allows principals to take an incremental approach to coaching individual teachers. Principal responses indicated that they believe their formative feedback is more meaningful to the teachers and more effective at improving teaching practices. Coaching is a relational strategy, and the majority of principals have better, or the same, relationships since using the *Framework*. Despite the overall positive impact on relationships, half of the principals felt that the impact of the *Framework* on instructional practices was dependent on the teacher's personality. Teachers who were less open to change or less receptive of the new observation tool would experience less improvement in their instructional practices.

Principals are now focused on instruction in new ways. This is well supported by their conversation about providing feedback and instructional coaching. Perhaps due to their experience with a less comprehensive observation model in the past, principals feel “focused” on instruction in new and deeper ways. Principals’ conversations with teachers are now focused on improving teaching practices. Overall, the relationships between principal and teacher are the same or better than before using the *Framework*.

Findings Related to Research Question 3

Research Question 3 states the following: To what extent, if any, do principals perceive that the use of the Danielson *Framework* is beneficial to the tenure and termination process?

Among summative decisions about job performance, all industries make important and consequential decisions when choosing to renew or terminate an employee’s employment. Principals also face these decisions, but no decision may be more consequential, across all industries, than the decision to award a teacher tenure. After months or years of formative feedback to help a teacher grow, the ultimate summative decision, a tenure decision, must occur. While formative assessment is intended to help teacher growth, summative assessment is intended to evaluate teacher practices (Scriven, 1967; Stufflebeam & Shinkfield, 2007).

Principals overwhelmingly attest that the *Framework* provides clear expectations for a teacher’s job performance (in instructional practices) and the necessary evidence to make a decision to recommend tenure for a teacher. Principals indicated that the expectations are clear in terms of both their specificity and their transparency. The specificity is similar to the principals’ belief that the *Framework* allows for targeted and

incremental growth; that is to say that expectations are clearly delineated. Principals discussed how they use the *Framework* to essentially publish their expectations of teacher's practices at varying levels of quality for their teachers. This allows for a mutual understanding between teachers and principals of the nature of different levels of teaching practices. The principals' discussion of this transparency echoes Danielson's own statement that the *Framework* provides a "shared understanding of teaching" (1996, 2007, p. 2)

Since using the *Framework*, principals believe the awarding of tenure has become more rigorous. The process for tenure used to be much more procedural; for example, after a certain number of years a teacher was granted tenure. Now, with the use of the *Framework*, principals are evaluating whether a teacher is proficient enough in various aspects of teaching to be deserving of tenure. Principals reflected that this evidence, or data, was helpful in both making the decision and substantiating the decision to the teachers and others.

With regard to summative evaluations (both contract renewals and tenure offers), principals are indicating that the expectations are clear and that they have also risen. This information does not directly align with principals' discussion about the credibility of ratings and relevance to all subjects and grade levels. In considering the evidence from both Research Question 1 and Research Question 3, principals are saying the following:

1. Their expectations for teacher instructional practices are clearer (but not clearer in all grade levels and all subjects)

2. They have more evidence to make summative decisions such as tenure (but they do not think their ratings are aligned with their principal colleagues' ratings)
3. Their expectations for tenure, and overall teacher performance, are more rigorous (but they feel the ratings are subjective based on the evaluator's views and expectations)

High expectations within the context of uncertainty and subjectivity are at odds with the literature on effective evaluations of personnel (Joint Committee on Standards for Educational Evaluation, 1988).

Findings Related to Research Question 4

Research Question 4 states the following: To what extent, if any, do principals perceive that the use of the Danielson *Framework* has improved their instructional leadership? The majority of principals felt that they have higher expectations for teachers' instructional practices since they began using the *Framework*, and an overall feeling that everyone can improve on something.

Half of the principals believe they are better able to lead instruction since using the *Framework*. In considering that nine principals indicated they are using the *Framework* for formative feedback (as discussed for Research Question 2) and only six believe they are better able to lead instruction, it seems that principals may not consider coaching, or individual, formative feedback to be new or improved since using the *Framework*.

Principals indicated that their instructional leadership improved both at the individual teacher level and at the school-wide level. Principals who indicated that their

instructional leadership has improved with individual teachers discussed the language provided by the *Framework*. Principals are using that specific language to enter conversations with teachers, serve as starting points in measuring a teacher's instructional practices, and helping provide support for a teacher's next steps. Principals who discussed improving their instructional leadership at the school-wide level discussed using the data from the evaluations to inform program choices, coaching candidates, and faculty-wide professional development.

Findings Related to Theoretical Frameworks

Scriven's Formative and Summative Evaluation

In "The Methodology of Evaluation," Michael Scriven introduced the terms *formative evaluation* and *summative evaluation* (1967). Scriven saw formative and summative as working, somewhat, together. As he described, the formative evaluation would occur during the development of the product or service, and the summative would occur after the development was completed. The final service or product, however, would have been formed, in part, by the information from the formative evaluation stage(s).

The principal participants in this study discussed their extensive use of the *Framework* for formative means. In terms of coaching and feedback, all principals discussed the *Framework* as a means of providing formative feedback for growth. Principals believe that the *Framework* provides the proper, specific language that allows them to provide targeted feedback that is meaningful for teachers. The *Framework* allows principals to focus on improving instruction with teachers in an incremental way that builds better relationships between principal and teacher.

In terms of the summative decisions (such as annual contract renewal and tenure), Scriven is unwavering that what is needed “are clear – but not simplistic – rules for deciding whether and how to reach justified conclusions” (Stufflebeam & Shinkfield, 2007, p. 381). If a rule is impossible, Scriven encourages evaluators to take other steps, including creating rubrics and/or calibrating evaluators. According to Scriven, the goal would be either through rule, rubric, calibration, or some other method by which different evaluators who are evaluating the same data would come to the same conclusion. If this is impossible, Scriven contends that a final judgment should not be offered but instead performance levels should be provided for separate aspects of the evaluation (Scriven, 1994).

The principal participants are not well aligned with Scriven’s statements on summative evaluation. Principals are mixed in their feelings about clarity. Principals do feel that the *Framework* provides clear expectations but also that principals’ judgments of those expectations are subjective and not aligned among colleagues. Only one in 12 principals believe that their understandings of the ratings are aligned with their principal colleagues; one third of principal participants feel that the *Framework* is not relevant to all subjects and grade levels; and zero principals have participated in calibration observations. In light of Scriven’s literature on summative assessment, it seems that final judgment would be difficult to support.

Fullan’s Leading Learning

Fullan’s theory does not strongly advocate for the use of an evaluation model to improve teacher practices and student learning. As he says, “You can use both—teacher evaluation and collaborative cultures—just get the order right” (Fullan, 2014, p. 77). This

study examined the use of the Danielson *Framework* with the lens of Fullan's learning leader.

Fullan's learning leader (or "leading learner" or "lead learner") is "one who models learning but also shapes the conditions for all to learn on a continuous basis" (Fullan, 2014, p. 9). The literature on effective principals focuses on collaborative cultures, learning communities, and capacity building (DuFour & Marzano, 2009; DuFour & Fullan, 2013; Fullan, 2010; Leithwood & Seashore-Louis, 2011; Hargreaves & Fullan, 2012). Success in providing highly effective teachers comes from systems that "help teachers work together in a focused way to use diagnostic student data linked to the improvement of instruction in order to get better results" (Fullan, 2014, p. 31).

Principal participants work often with individual teachers. Principals discussed at length the individual instructional coaching and use of the *Framework* to identify individual teacher's strengths and weaknesses. The principals' discussion of individual teacher's actions (such as seeking out the principal for conversations about instruction) and the receptiveness of teachers to feedback (based on their personality) further supports the use of the *Framework* to improve teacher practices at the individual teacher level. Despite most principals discussing their work with individual teachers, a few (3) discussed their use of the *Framework* to address larger, macro, instructional issues with the larger faculty. Four principals used the *Framework* to inform their design and implementation of professional development (for groups larger than a single teacher).

Seven of the 12 principals did discuss their use of the *Framework* to inform their work around building capacity, collaborative cultures, or learning communities (directly aligned with Fullan's learning leader). These efforts were largely centered on building

collaborative cultures and did not really include the combination of two or three attributes of the learning leader.

Recommendations for Policy

The importance that teachers and principals play in the achievement of students is well documented, and therefore policymakers can spare no effort in improving the instructional impact of teachers in the classroom.

1. Include a requirement in the policy statement for ongoing training of principals and other evaluators. Principals do not feel properly prepared for evaluating teachers with the Danielson *Framework*. Many principals found the model to be foreign in concept and practice, despite their enthusiasm for its usefulness, which certainly led to problems. The principals had difficulty using the model in all grade levels and subjects and have a lack of rater reliability. Recognizing that new principals, and other evaluators, need to be trained differently than the ongoing support needed for veteran evaluators, policymakers should work to create a comprehensive, systemic, and ongoing professional development series for observation and evaluation.
2. Determine a consistent policy for the evaluation of special teachers and early childhood teachers. Principals indicated that the *Framework* as written does not help in evaluating and providing feedback in special subjects (Physical Education, World Languages, Art, Music, etc.) and early childhood classrooms. A lack of a clear manner for evaluating these teachers will lead to similar consequences as the lack of valid ratings (cited in Recommendation for Policy #3).

3. Expand the tenure process to account for more data points. Principals indicated the reliability of their ratings were not aligned. Policymakers should take note of this and expand the data points used in determining teacher tenure (and possibly contract renewals as well) to ensure that tenure decisions are made with the highest degree of reliability across classrooms and schools.

Recommendations for Practice

Principals can effect great change within their schools and districts. Making a few shifts in practice could significantly improve their instructional leadership and therefore teachers' instructional practices and student achievement.

1. Train principals to ensure greater validity in ratings. As has been previously stated, only one in 12 principals believed the ratings to be aligned with their colleagues. The hallmark of an evaluation system must be that it is valid across classrooms and schools. If the principals are indicating that they are not in alignment with one another, policymakers must take actions to bring the principals into alignment. The lack of alignment between principals' ratings could be greatly mitigated by the use of calibration observations as a main component of the training. The principals indicated that they already come together for small and large sized principal meetings. These meetings, if not already held, could be held at schools and a certain amount of time could be used for calibration. To calibrate, principals would only need to observe a lesson as a group for a period of time and then debrief the lesson (discussing their ratings and why they would give that rating).

2. Develop school-based cultures focused on personal, instructional growth built on trust. Principals discussed challenges with certain types of personalities among teachers. These personalities were often focused on a fixed mindset, and principals should utilize strategies and systems that help develop cultures of growth for the adults as well as the students. The principals often converse with teachers about instruction. Using conversation protocols will help with two areas: (a) ensuring that these conversations are always actionable for teachers and (b) limiting teachers' negative reactions by following a common approach. These conversations matter, and they must be considered as vital a part of the evaluation process as the ratings and the written report.

Principals discussed having pre- and post-conferences with teachers, and both conversations need to be meaningful to teachers and provide actionable next steps for teachers. Principals indicated they sometimes encounter teachers with difficult personalities that therefore limits the impact of the *Framework* on that teacher's growth. Using a conversation protocol will help in establishing norms of the conversation that the teacher understands to be universal to all teachers, therefore helping to remove personal conflict and build trust. In order to maximize these conversations, the principals should use meeting, or conversation, protocols that are research based. A growth mindset is important, but building a culture of trust with teachers must be a top priority of a principal.

3. Allocate district financial resources to target professional development and data tracking. School leaders and policymakers could look for instructional

trends across the district to make determinations about professional development and other instructional supports. Currently, the teacher observation information is located as separate, usually paper, documents in each school. Policymakers should consider migrating the data to a digital software that would allow for ongoing monitoring of instructional strengths and weaknesses. Principals indicated that they had generally not used observation/evaluation data in their design and implementation of professional development. Policymakers could use this information in designing professional development for teachers and for monitoring and providing supports for principals' evaluation practices.

Recommendations for Future Research

The focus on teacher effectiveness appears to remain in the minds of policymakers and parents alike. Due to this focus, the continued study of teacher evaluation models is essential because teacher evaluation models are likely the main tool used to turn that focus on teacher improvement into action. The more that is understood about teacher evaluation the more policymakers can make informed decisions.

1. Expand the number of participating principals. Given that this study was a small study seeking to understand a phenomenon, the study could be expanded to more principals in this district or other districts.
2. Conduct a quantitative study of principals' perceptions. This study has generated some conclusions and themes that could be further studied in a quantitative or mixed-method study. The Seven Fountains School District is a sister district of eight other districts, all of which use the same evaluation

model. A study could be conducted that captures a greater sample of principals from those sister districts.

3. Investigate the impacts of different training and education programs used by principals. Every principal has a different background which shapes his or her understanding of effective teaching practices. All of the principal participants in this study had master degrees; but when, what type, and whether there was an evaluation focus could shape their practice of evaluation. Studying different principal training and educational programs could provide strategies for how to best prepare principals for evaluating teacher performance.
4. Explore effective feedback for teachers. Providing actionable and meaningful feedback to teachers is imperative to a functioning evaluation system. The Danielson model provides the ability to evaluate teacher practices, but it does not provide the next steps for improvement. A deeper understanding of the nature of effective feedback for teachers would be helpful for principals in leading instruction and providing formative feedback.
5. Review best practices in the process of evaluating teachers. The Danielson model does not lay out effective systems and structures for principals to follow in the process of evaluating teachers. Publishing best practices and approaches to teacher evaluation could be beneficial to principals who are struggling to accomplish everything that a particular evaluation model requires.
6. Inquire about the experience of teachers who were not granted tenure or were not renewed for another year of teaching. By uncovering themes and

experiences of “unsuccessful” teachers, principals could gain valuable insights into why some teachers were able to grow and some were not. Principals could learn where the barriers to growth occur and potentially discover some strategies to better set teachers up for success.

7. Explore the impact of a new teacher evaluation model on student achievement. This study was conducted in the early years (the third year) of the use of the *Framework*. In a few more years, it could and should become possible to evaluate any effects on student achievement.
8. Probe the impact of the new teacher evaluation model on teacher performance indicators, including *Framework* ratings (although those may not be highly credible), contract renewal rates, and tenure rates. Since this study was conducted during the third year of the *Framework*’s implementation, in a few more years there could be sizable data points connected to questions about impact on teacher performance.

References

- American Psychological Association. (2017, January 1). *Ethical principles of psychologists and code of conduct*. Retrieved from <http://www.apa.org/ethics/code/index.aspx?item=4>
- Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., & Shepard, L. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute.
- Barge, J. D. (2012). *Overview to the 2012 TKES/LKES Pilot evaluation report* . Atlanta, GA: Department of Education, State of Georgia.
- Barton, P., Darling-Hammond, L., Haertel, E., Baker, E., Ladd, H., Linn, R., . . . Shepard, L. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute.
- The Bill and Melinda Gates Foundation. (2010). *K-12 Education*. Retrieved from http://k12education.gatesfoundation.org/teacher-supports/teacher-development/measuring-effective-teaching/?gclid=CLqV_6P1zNICFZWCSwodYpoM5Q
- The Bill & Melinda Gates Foundation. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle, WA: Author.
- Bogdan, R. C., & Biklen, S. K. (2003). *Qualitative research for education: An introduction to theories and methods* (4th ed.). Boston, MA: Pearson.
- Brandt, C., Mathers, C., Olivia, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest*

- region (Issues and Answers Report, REL 2007- No. 030)*. Washington, DC: Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, REL Midwest.
- Carnegie Forum on Education and the Economy. (1986). *A Nation Prepared: Teachers for the 21st Century. The Report of the Task Force on Teaching as a Profession*. New York, NY: Author.
- Cogan, M. (1964). *Clinical supervision by groups*. Cedar Falls, IA: The Association for Student Teaching.
- Coughlan, M., & Cronin, P. (2007). Step-by-step guide to critiquing research. *British Journal of Nursing*, 36(12), 738-744.
- Creswell, J. W. (2002a). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. Columbus, OH: Merrill Prentice Hall.
- Creswell, J. W. (2002b). *Research design: Qualitative, quantitative, and mixed method approaches*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage.
- Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory Into Practice*, 39(3).
- Cronbach, L. J. (1963). Course improvements through evaluation. *Teachers College Record*, 64, 672-683.
- Cubberley, E. (1916). *Public school administration: A statement of the fundamental principles underlying the organization and administration of public education*. Boston, MA: Houghton Mifflin.

- D'Amico, D. (2011, January 13). NJEA plans to fight Gov. Christie over plan to eliminate teacher tenure. *The Press of Atlantic City*.
- Danielson, C. (1996, 2007). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C. (2013). *The framework for teaching: Evaluation instrument*. Princeton, NJ: Danielson Group.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C., & McGreal, T. L. (2014). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- The Danielson Group. (2013). *The Framework*. Retrieved from <https://www.danielsongroup.org/framework/>
- Darawsheh, W. (2014). Reflexivity in research: Promoting rigour, reliability and validity in qualitative research. *International Journal of Therapy & Rehabilitation*, 21(12), 560-568.
- Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching*. New York, NY: National Commission on Teaching and America's Future.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis Archives*, 8(1).

- Darling-Hammond, L., & Ducommun, C. (2010). *Recognizing and developing effective teaching: What policy makers should know and do*. Washington, DC: The Partnership for Teacher Quality.
- Dee, T., & Wyckoff, J. (2013). *Incentives, selection, and teacher performance: Evidence from impact*. Cambridge, MA: National Bureau of Economic Research.
- DuFour, R., & Fullan, M. (2013). *Cultures built to last: Systemic PLCs at work (Help lead and sustain your school improvement process)*. Bloomington, IN: Solution Tree.
- DuFour, R., & Marzano, R. J. (2009, February). High-level strategies for principal leadership. *Educational Leadership*, 66(6), 62-68.
- Editorial Projects in Education Research Center. (2015, September 3). Issues A-Z: Teacher evaluation: An issue overview. Retrieved from <http://www.edweek.org/ew/section/multimedia/teacher-performance-evaluation-issue-overview.html/>
- Eggen, P., & Kauchak, D. (1996). *Strategies for teachers: Teaching content and thinking skills*. Needham Heights, MA: Allyn & Bacon.
- Eisner, E. W. (1975). *The perceptive eye: Toward the reformation of educational evaluation*. Stanford, CA: Stanford Evaluation Consortium.
- Ellet, C., & Teddlie, C. (2003, March). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, 17(1), 101-128.

- Eskro-Clemetson, H. (2000). *Comparison of perceptions of elementary teachers about teacher evaluation procedures in two different school districts* (Unpublished master's thesis). Wayne State University, Detroit, MI.
- Fullan, M. (2010, March/April). The awesome power of the principal. *Principal*, 10-15.
- Fullan, M. (2014). *The principal: Three keys to maximizing impact*. San Francisco, CA: John Wiley & Sons.
- Gallagher, H. (2004). Vaughner Elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education*, 79(4), 79-107.
- Gallup. (2009, September 3). Labor unions see sharp slide in U.S. public support. *Gallup News*.
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18, 519-521.
- Glesne, C. (1999). *Becoming qualitative researchers: An introduction* (2nd ed.). Boston, MA: Pearson.
- Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, 8(4), 597-606.
- Goldhammer, R. (1969). *Clinical supervision: Special methods for the supervision of teachers*. New York, NY: Holt, Rinehart, and Winston.
- Goldstein, D. (2014). *The teacher wars: A history of America's most embattled profession*. New York, NY: Random House.

- Haefele, D. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education* , 7.
- Hammond, R. L. (1967). *Evaluation at the local level*. Address to the Miller Committee for the National Study of ESEA Title III, Washington, DC.
- Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. New York, NY: Teachers' College Press.
- Harris, B. M. (1965). Strategies for instructional change: Promising ideas and perplexing problems. In J. Roth and R. Leeper (Eds.), *The supervisor: Agent for change in teaching* (p. 129). Washington DC: Association for Supervision and Curriculum Development.
- Heneman, H. G., Milanowski, A., Kimball, S. M., & Odden, A. (2006, May). *Standards-based teacher evaluation as a foundation for knowledge- and skill-based pay*. Philadelphia, PA: Consortium for Policy Research in Education. Retrieved from <http://www.cpre.org/standards-based-teacher-evaluation-foundation-knowledge-and-skill-based-pay>
- Hogan, R. L. (2007). The historical development of program evaluation: Exploring the past and present. *The Online Journal of Workforce Education and Development*, 2(4), 1-14. Retrieved from <http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1056&context=ojwed>
- Joint Committee on Standards for Educational Evaluation. (1988). *The personnel evaluation standards*. Thousand Oaks, CA: Corwin Press.

- Keesler, V. A., & Howe, C. (2012). *Understanding educator evaluations in Michigan: Results from 1 year of implementation*. Lansing, MI: Michigan Department of Education.
- Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, 79(4), 54-78.
- Klein, A. (2015). K-12 law's legacy blend of idealism, policy tensions. *Education Week*, 34(26), 1-21.
- Kratz, H. (1896). Characteristics of the best teachers as recognized by children. *The Pedagogical Seminary*, 3(3).
- Leithwood, K., & Seashore-Louis, K. (2011). *Linking leadership to student learning*. San Francisco, CA: Jossey-Bass.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. New York, NY: Wallace Foundation.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverley Hills, CA: Sage.
- Lyons, G. (1980, October). Why teachers can't teach. *Phi Delta Kappan*, 62(2).
- Marzano, R., Pickering, D., & Pollock, E. (2011). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McMillan, J. H., & Schumacher, S. (1997). *Research in education: A conceptual introduction* (4th ed.). New York, NY: Longman.
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

- Metfessel, N. S., & Michael, W. B. (1967). A paradigm involving multiple criterion measures for the evaluation of the effectiveness of school programs .
Educational and Psychological Measurement, 27, 931-943.
- Milanowski, A. (2004). The relationship between teacher performance evaluation schools and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, 79(4), 33-53.
- Miles, J., & Gilbert, P. (2005). *A handbook of research methods for clinical and health psychology*. New York, NY: Oxford University Press .
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Government Printing Office.
- National Council on Teacher Quality. (2011). *State of the states: Trends and early lessons on teacher evaluation and effectiveness policies*. Washington, DC: National Council on Teacher Quality .
- National Council on Teacher Quality. (2016). *State-by-state evaluation timeline briefs*. Washington, DC: National Council on Teacher Quality .
- Organization for Economic Cooperation and Development. (2016). *PISA 2015 Results in Focus* . OECD. Retrieved from www.oecd.org/pisa
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Saint Paul, MN: Sage.

- Peterson, C. H. (1982). *A century's growth in teacher evaluation in the United States* . New York, NY: Vantage Press.
- Pink, D. (2009). *Drive: The surprising truth about what motivates us*. New York, NY: Riverhead Books.
- Popham, W. J. (1971). *Criterion-referenced measurement* . Upper Sadle River , NJ: Educational Technology Publications .
- Provus, M. (1969). *Discrepancy evaluation model*. Pittsburgh, PA: Pittsburgh Public Schools.
- Reckhow, S. (2013). *Follow the money: How foundation dollars change public school politics* . New York, NY: Oxford University Press .
- Roberts, P., Priest, H., & Traynor, M. (2006). Reliability and validity in research. *Nursing Standard*, 20(44), 41-45.
- Rogers, C. H., & Weems, D. M. (2010). Are U.S. teachers making the grade? *Management in Education*, 24(1).
- Russakoff, D. (2015). *The prize: Who's in charge of America's schools?* . New York, NY: Houghton Mifflin Harcourt.
- Saltman, K. (2010). *The gift of education: Public education and venture philanthropy*. New York, NY: Palgrave MacMillan.
- Sanjari, M., Bahramnezhad, F., Khoshnava Fomani, F., Shoghi, M., & Ali Cheraghi, M. (2014). Ethical challenges of researchers in qualitative studies: The necessity to develop a specific guideline. *Journal Of Medical Ethics & History Of Medicine*, 7(14), 1-6.

- Sawchuk, S. (2013, February 6). High ratings for teachers are still seen: Changes to evaluations systems yield only subtle differences. *Education Week*, 1.
- Scott, J. (2006). The politics of venture philanthropy in charter school policy and advocacy . *Education Policy*, 23(1), 106-136.
- Scriven, M. S. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, and M. Scriven (Eds.), *Perspectives of curriculum evaluation* . Skokie, IL: Rand McNally.
- Scriven, M. S. (1974). Pros and cons about goal-free evaluation. *Evaluation Comment*, 3, 1-4.
- Scriven, M. S. (1994). The final synthesis. *Evaluation Practice*, 15(3), 367-382.
- Shough, L. L. (2010, May). *Standards-based teacher evaluation: A professional development tool for improving instruction* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3407127)
- Spencer, L., Ritchie, J., & O'Connor, W. (n.d.). Analysis: Practices, principles and processes. In J. Ritchie, & J. Louis (Eds.), *Qualitative research practice: A guide for social scientist students and researchers* (pp. 199-218). London: Sage.
- Stake, R. E. (1967). The countenance of educational evaluation. *Teachers College Record*, 68, 523-540.
- Strauss, V. (2016, December 6). Why Americans should not panic about international test results. *The Washington Post*.
- Stufflebeam, D. L. (1967). The use and abuse of evaluation in Title III. *Theory into practice*, 6, 126-133.

- Stufflebeam, D. L. (1971). The relevance of the CIPP evaluation model for educational accountability. *Journal of Research and Development in Education*, 5(1), 19-25.
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.
- Surowiecki, J. (2011, January 17). State of the unions. *The New Yorker*.
- Taylor, E. S., & Tyler, J. H. (2012, Fall). Can teacher evaluation improve teaching?: Evidence in the systematic growth in the effectiveness of midcareer teachers. *Education Next*, 12(4) Retrieved from <http://educationnext.org/can-teacher-evaluation-improve-teaching/>
- Taylor, F. (1911). *Scientific Management*. New York and London: Harper Brothers' Publishers.
- Tienken, C. (2013, April-June). Conclusions from PISA and TIMSS testing. *Kappa Delta Pi Record*, 49, 56-58.
- Tucker, P. D., & Stronge, J. H. (2005). *Linking teacher evaluation and student learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. Chicago, IL: University of Chicago Press.
- Tyler, R. W. (1967). Changing concepts of educational evaluation. In R. E. Stake (Ed.), *Perspectives of curriculum evaluation* (pp. 13-28). Skokie, IL: Rand McNally.
- U.S. Department of Education. (2004, March). New No Child Left Behind flexibility: Highly qualified teachers. Retrieved from <https://www2.ed.gov/nclb/methods/teachers/hqtflexibility.html>

- U.S. Department of Education. (2009). Race To the Top resources. Retrieved from <http://www2.ed.gov/programs/racetothetop/phase1-resources.html>
- U.S. Department of Education. (2009, November). *Race to the Top program executive summary*. Washington, DC: Author.
- U.S. Department of Education. (2010, January). Race to the Top Fund scoring rubric. Retrieved from <https://www2.ed.gov/programs/racetothetop/scoringrubric.pdf>
- U.S. Department of Education. (n.d.). Every Student Succeeds Act. Elementary and Secondary Education Act as amended by the Every Student Succeeds Act. Retrieved from <https://www.ed.gov/essa?src=rn>
- Walsh, K., Joseph, N., Lakis, K., & Lubell, S. (2017). *Running in place: How new teacher evaluations fail to live up to promises*. Washington, DC: National Council on Teacher Quality.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The Widget Effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York, NY: The New Teacher Project.
- Weller, R. H. (1971). *Verbal communication in instructional supervision: An observational system for research and study of clinical supervision in groups*. New York, NY: Teachers College Press.
- Wetzel, W. (1929). Scientific supervision and curriculum building. *The School Review*, 37(2), 117-123.
- The White House. (n.d.). President Barack Obama. Retrieved from <https://obamawhitehouse.archives.gov/the-press-office/fact-sheet-race-top>

Wiggins, G., & McTigue, J. (1998). *Understanding by design*. Alexandria, VA:
Association for Supervision and Curriculum Development.

Appendix A



OFFICE OF THE SUPERINTENDENT OF SCHOOLS ARCHDIOCESE OF NEW YORK

1011 First Avenue, New York, New York 10022 Phone: (212) 371-1000 Fax: (212) 758-3018

August 24, 2017

Seton Hall University
400 South Orange Avenue
South Orange, NJ 07079

Attn: Institutional Review Board

This letter serves to confirm approval for Stephen Hernon to conduct research in the [REDACTED] Region in the Archdiocese of New York for his dissertation: An Exploration of Principal Perceptions of the Danielson Framework for Teaching in Providing Feedback and Improving Instructional Leadership.

Under supervision of [REDACTED], Regional Superintendent, Catholic School Region of [REDACTED], Archdiocese of New York, Seton Hall agrees to comport with all archdiocesan policies and procedures and share with us a copy of the final document.

Sincerely,

Timothy J. McNiff, Ed.D.
Superintendent of Schools
Archdiocese of New York

CC: [REDACTED], Regional Superintendent, Catholic School Region of [REDACTED]
[REDACTED], Archdiocese of New York

Appendix B



January 11, 2018

Steve Hernon
[REDACTED]
[REDACTED]

Dear Mr. Hernon,

The Seton Hall University Institutional Review Board has reviewed the information you have submitted addressing the concerns for your proposal entitled "An Exploration of Principal Perceptions of the Danielson Framework for Teaching in Providing Feedback and Improving Instructional Leadership". Your research protocol is hereby approved as revised through expedited review. 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 form.

Also enclosed is the \$1.00 bill submitted in your envelope.

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 your 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.

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

cc: Dr. Barbara Strobert

Office of Institutional Review Board
Presidents Hall • 400 South Orange Avenue • South Orange, NJ 07079 • Tel: 973 313 6314 • Fax: 973 275 2361 • osir@shu.edu

Appendix C

