

Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

12-2018

An Investigation of Educators' Perceptions of the Influence of a Teacher Evaluation System on Student Learning

Bryon Nielsen
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Education Commons](#)

Recommended Citation

Nielsen, Bryon, "An Investigation of Educators' Perceptions of the Influence of a Teacher Evaluation System on Student Learning" (2018). *All Graduate Theses and Dissertations*. 7327.
<https://digitalcommons.usu.edu/etd/7327>

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



AN INVESTIGATION OF EDUCATORS' PERCEPTIONS OF THE INFLUENCE
OF A TEACHER EVALUATION SYSTEM ON STUDENT LEARNING

by

Bryon Nielsen

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF EDUCATION

in

Education
(Curriculum and Instruction)

Approved:

Courtney Stewart, Ph.D.
Major Professor

Michael Freeman, Ph.D.
Committee Member

Tyson Sorensen, Ph.D.
Committee Member

Sylvia Read, Ph.D.
Committee Member

Susan Turner, Ph.D.
Committee Member

Laurens H. Smith, Ph.D.
Interim Vice President for Research and
Interim Dean of the School of Graduate
Studies

UTAH STATE UNIVERSITY
Logan, Utah

2018

Copyright © Bryon Nielsen 2018

All Rights Reserved

ABSTRACT

An Investigation of Educators' Perceptions of the Influence of a Teacher
Evaluation System on Student Learning

by

Bryon Nielsen, Doctor of Education
Utah State University, 2018

Major Professor: Courtney Stewart, Ph.D.
Department: Teacher Education and Leadership

In this investigation, the perceptions, attitudes, and beliefs of teachers and administrators regarding a teacher evaluation system were examined. The purpose of the investigation was to gain greater understanding of how the perceptions, attitudes, and beliefs of teachers and administrators influence the implementation and efficacy of the evaluation system. Knowledge gained from this study may be helpful in developing strategies for implementation of evaluation systems in ways that promote the efficacy of the tool. Results of this study may influence related research concerning evaluation and supervision and the extent to which comparable evaluation systems serve their intended purpose. A qualitative analysis of both documents and interview responses served to identify educator perceptions regarding the evaluation process and how those perceptions influence the implementation of the evaluation tool. Strategies are recommended that promote effective implementation of the evaluation system in ways that increase the likelihood that the intended purposes of evaluation will be realized.

(223 pages)

PUBLIC ABSTRACT

An Investigation of Educators' Perceptions of the Influence of a
Teacher Evaluation System on Student Learning

Bryon Nielsen

In this study, the researcher explored the views and opinions held by teachers and administrators toward a teacher evaluation system. The purpose was to learn how educators' opinions and views might impact how school leaders administer the evaluation system and how teachers receive it. Knowledge gained from this study may provide understanding for district and school-level leaders as they develop training and implementation strategies. Teachers and administrators were interviewed, and district-provided documents were reviewed, to gather data that was analyzed using qualitative research methods. Based on the data analysis, recommendations are made regarding implementation strategies that may improve the evaluation process.

ACKNOWLEDGMENTS

I would like to acknowledge the many faculty in the department of Teacher Education and Leadership who have assisted me in this journey. Especially those who served as part of my committee, for their willingness to let me try until I got it, and for their patience as I tried to figure things out. I enjoyed every class and the opportunity to learn from and be tutored by outstanding teachers and professors.

I am indebted to those wonderful individuals who made up our cohort who also helped me along the way. Without them, the journey would have been far more difficult and much less enjoyable. I will always be grateful to them for their kindness, friendship, and example of scholarship and achievement.

I am thankful for the support of my wonderful wife and family who encouraged me and supported me through the extra time away and when I may have been tired and discouraged. They are the best people I know. They have always been the impetus for any improvement in my life and what I have endeavored to do to be a better person.

Bryon Nielsen

CONTENTS

	Page
ABSTRACT.....	iii
PUBLIC ABSTRACT	iv
ACKNOWLEDGMENTS	v
LIST OF TABLES.....	viii
LIST OF FIGURES	ix
CHAPTER	
I. INTRODUCTION.....	1
Historical Background.....	2
Purpose	4
Summary	15
II. LITERATURE REVIEW	17
Theoretical Framework	18
Historical Context.....	24
Characteristics of Contemporary Evaluations.....	33
Role of the Principal.....	42
Current Issues	43
Summary	65
III. METHODOLOGY	68
Design.....	68
Research Questions	69
Data Collection.....	69
Summary	83
IV. RESULTS.....	84
Summary of Data.....	87
Data Reduction Process.....	90
Findings Related to Research Questions	91

	Page
Analysis	126
V. CONCLUSIONS	152
Findings	152
Recommendations	170
Limitations and Further Research	183
Conclusion	185
REFERENCES	188
APPENDICES	199
Appendix A: Permission Letter from Davis School District to Use Evaluation Documents	200
Appendix B: Interview Protocol.....	202
Appendix C: Informed Consent Form.....	204
Appendix D: PLAN Rubric	208
Appendix E: Administrator and Teacher Interview Schema, Document Review Schema.....	210
Appendix F: Links to Effective Utah Teaching Standards and <i>Evaluate Davis</i> Manuals	212
CURRICULUM VITAE.....	214

LIST OF TABLES

Table	Page
1. Differences Between the Previous System and Evaluate Davis	10
2. Coding Schema Teacher Focus Group	79
3. Teacher Demographic Information.....	90
4. Administrator Demographic Information	90

LIST OF FIGURES

Figure	Page
1. Evaluation perception cycle.....	48
2. Demographic information form.....	72

CHAPTER I

INTRODUCTION

Federal and state attention to school and teacher accountability for student learning has increased dramatically in the last decade and with it has come growing importance for evaluating teachers' impact on student achievement. Recent research that documents the impact of teacher quality on student learning (Akiba, LeTendre, & Scribner, 2007) has merged with an increasing understanding that teacher evaluation has been ineffective (Donaldson, 2009; Weisberg, Sexton, Mulhern, & Keeling, 2009). Current teacher evaluation reform efforts have been the outcome (Donaldson & Papay, 2015). Despite a lack of evidence that evaluation reforms are improving teacher performance (Kraft & Gilmour, 2017), states and districts across the U.S. are making changes to their teacher evaluation procedures with the expectation that the outcome will be better teaching and improved student achievement (Donaldson & Papay, 2015; Goe, Bell, & Little, 2008; Kimball & Milanowski, 2009; Popham, 2013; Steinberg & Sartain, 2015).

However, many teachers, principals, and researchers are suspicious of the changes and unsure of their usefulness in terms of improving teaching (Goodwin & Hein, 2016; Kraft & Gilmour, 2016). In this study, I examined principals' attitudes about evaluation and how their attitudes may impact, for good or bad, how they implement the evaluation within their school. Likewise, I also studied teachers' attitudes about the evaluation process with respect to how their attitudes influence how they accept and react to the feedback they are given during the evaluation process (Kimball & Milanowski, 2009).

Historical Background

Supervising and evaluating teachers has been part of schooling in the United States since colonial times. Then teachers were considered servants of the community, so local governments used clergy and supervisory committees to hire and fire teachers and monitor instructional quality (Sullivan & Glanz, 2009) although little was done to improve teaching (Tucker & Stronge, 2005). Until the mid-19th century, supervision of teachers remained somewhat static, focused on maintaining, rather than improving, existing standards of teaching (Mielke, 2012).

As school systems became more complex during the Industrial Revolution in the 1800s, the construct of effective teaching (what teacher behaviors constitute successful teaching) became more important. The job of evaluating and supervising teachers became more than simply monitoring their behavior; teachers were expected to know how to teach and to improve their teaching skills (Marzano, Frontier, & Livingston, 2011). With the shift from an agricultural society to an industrial came a growing need for a more educated workforce. As well, more families began living in concentrated areas, creating a greater demand for schools and teachers. The responsibility for supervising and evaluating teachers moved from community leaders such as clergy or elected officials to superintendents who were hired specifically to manage and supervise schools and teachers. As the task grew beyond superintendents' capacity, principals of individual schools were hired and given the role of supervising teachers. Even until the mid-1900s, the purpose of evaluation was largely supervisory, used to make personnel and compensation decisions, and assure quality, but not for teacher improvement (Popham,

1988).

As American society has grown more complex, our public schools' needs have transformed as well, but at an increasing rate. Scientific models of management, which focused on the purpose of education and how teachers should teach, were introduced early in the Twentieth Century. Formative Clinical Supervision Models of teacher evaluation were introduced during the 1960s and the 1970s (Marzano et al., 2011) as policy makers began to look at improving public schools by improving teaching. During the 1980s, more technical or didactic models of teaching emerged from the effective teaching literature (Hunter, 1986) and with that came models of evaluation and supervision that were also increasingly technical. Teachers were told to implement specific strategies and programs that were expected to raise the level of learning and put America back in the education race. The basic elements of the mostly-formative clinical supervision models and Hunter's seven-step model persist in the standards-based models of the current decade (Marchiando, 2013). Today, policy makers at the state and national level have an expectation that better teacher evaluations will improve the nation's public schools and make high levels of learning accessible to all students (Donaldson, 2016). In the last decade, Federal initiatives, such as Race to the Top and the Elementary and Secondary Education Act (ESEA) Flexibility Program, mandated that states revamp their teacher evaluation systems, requiring them to include elements such as increased classroom observations, parent and student surveys, and measures of student academic growth with the goal of improving the quality of teaching in our nation's public schools (Goe, 2013; Goe et al., 2008; Popham, 2013; Towe, 2012). Recently, the Every Student

Succeeds Act of 2015 (ESSA), which replaced No Child Left Behind, relaxed some of the mandates placed on states to include specific components in their teacher evaluation systems (Sawchuck, 2016). Still, Donaldson (2016) reported that forty-six states have overhauled their teacher evaluation systems since 2009, incorporating multiple rating categories such as student growth measures, classroom observations, and parent and student surveys.

Similar to many school districts across the country, the present evaluation system in Davis School District (DSD), in Davis County, Utah, titled *Evaluate Davis*, incorporates current teacher evaluation research (Danielson, 2011; Goe et al., 2008; Kane, Taylor, Tyler, & Wooten, 2011, Marzano & Toth, 2013; Mathers & Olivia, 2008; Pallas, 2010). The focus of new standards-based evaluations, such as *Evaluate Davis*, is on teacher improvement. Likewise, there is an expectation of *Evaluate Davis* that as teaching improves, student-learning gains will increase. Many new research-based evaluations are similar to *Evaluate Davis* and include the following research-based components: (a) multiple data sources, such as several classroom observations, parent and student surveys, and student achievement data; (b) self-evaluation; (c) goal setting; and (d) aligned professional development (Danielson, 2011; Danielson & McGreal, 2000; Darling-Hammond, 2012; Donaldson, 2016).

Purpose

This research was designed to examine teachers' and administrators' experience with the new teacher evaluation tool, *Evaluate Davis*, considering specifically how

teachers' and principals' perceptions, attitudes, and beliefs about evaluation influence implementation of the evaluation process.

Problem Statement

Current literature indicates a substantial disparity between the expectations of researchers and policy makers on one hand and the perceptions and attitudes of practitioners on the other regarding the influence of teacher evaluation on teacher quality and student learning. New standards-based teacher evaluation systems are predicted to be an important factor in improving teaching and raising student achievement, and yet many principals and teachers harbor negative perceptions, beliefs, and attitudes of teacher evaluation that may hinder its impact (Goe, 2013; Goe, et al., 2008; Popham, 2013; Towe, 2012). I examined several factors contributing to the disparity identified above: (a) recommendations of current research; (b) the expectations of policy makers; (c) the attitudes, behaviors, and perceptions of administrators and teachers; (d) principals' will and skill regarding evaluation; (e) principals' lack of time; and (f) the difficulty in linking current teacher evaluation practices to improved teaching and increased student learning.

Current research. Recent research suggests that more rigorous teacher evaluation practices may be an effective means for improving teaching (Danielson & McGreal, 2000; Marzano & Toth, 2013; Towe, 2012; Tucker & Stronge, 2005). Researchers (Donaldson, 2009) have looked increasingly to evaluation as the path to improving teacher effectiveness. Halverson, Kelley, and Kimball (2004) reported that the consensus among researchers and policymakers is that improved evaluations will improve teaching, and in turn, better teaching will raise student achievement.

Expectations. The recent accountability movement has added momentum to the current trend in the U.S. to rely on evaluation systems to improve teaching, improve schools, and increase student achievement (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; Donaldson, 2016). Kimball & Milanowski (2009) reported that the current climate of accountability in American education encourages the use of evaluation to make high-stakes decisions regarding teacher salary and retention and to improve teacher quality. Policy makers at the state and national level have an expectation that better, more rigorous evaluations will improve our nation's public schools and make high levels of learning accessible to all students (Goe, 2013; Goe et al., 2008; Popham, 2013; Towe, 2012). Recent Federal initiatives, such as Race to the Top and the Elementary and Secondary Education Act (ESEA) Flexibility Act, have encouraged education reform efforts focusing on increasing the rigor of teacher evaluation methods with the expectation that the result will be improved teacher practice and increased academic achievement among all students (Steinberg & Sartain, 2015).

Poor perceptions. According to Halverson et al. (2004) teachers' and principals' perceptions of traditional teacher evaluation have been negative, viewed as perfunctory, end-of-year rituals that have little or no bearing on the quality of teaching or the effectiveness of teachers (Pallas, 2010; Peterson, 2004; Popham; 2013). Recent research revealing that traditional teacher evaluation systems have been ineffective in both summative and formative aspects, unsuccessful in discriminating between poor and effective teachers, and lacking influence in developing teacher effectiveness (Donaldson, 2009; Weisberg et al., 2009), has contributed to lingering negative attitudes about

evaluation processes. Goodwin and Hein (2016) reported that some have “questioned whether the tremendous effort to observe teachers in classrooms, given its weak overall correlation to achievement, was worth it” (p. 83). Schmoker (2012) complained that “...we are rushing into a premature, ill-conceived innovation without any solid evidence that it promotes better teaching” (p. 20).

Skill and will. The burden of teacher evaluation rests primarily with principals who determine the usefulness of the system through their implementation of the system’s components (Steinberg & Sartain, 2015). Principals, however, often lack the skills of gathering and interpreting data and following it up with quality feedback to provide meaningful instructional guidance (Goe, 2013; Donaldson, 2010) which deprives teachers of the component of the evaluation process they value most (Donaldson, 2016) further diminishing their perception of the system. According to Donaldson (2010) principals who are lacking in training are more likely themselves to treat the process as perfunctory, give meaningless feedback, and rate all teachers as above average most of the time.

For many principals, the lack of will to both implement all of the components of the evaluation system and give less-than perfect ratings and feedback (Peterson, 2004; Tucker & Stronge, 2005) also makes the evaluation less effective in discriminating between teachers and identifying individual teachers’ strengths and weaknesses (Kimball & Milanowski, 2009). Absent the ability to effectively discriminate between various levels of teaching quality, teachers and administrators default to seeing the evaluation as a validation and reward rather than a formative tool for improvement (Danielson, 2001), further undermining the efficacy of the evaluation.

Time. Administrators have so many other responsibilities taking their time that they have traditionally spent little time in classrooms, providing meaningful feedback, or engaging in conversations with teachers regarding targeted professional development (Donaldson, 2016). Consequently, teachers, who value observations, feedback, and formative conversations, have found little developmental worth in conventional evaluations (Zatynski, 2012).

However, new standards-based teacher evaluations, such as *Evaluate Davis* (Davis School District, 2014) and the Chicago Public Schools teacher evaluation system Steinberg and Sartain (2015) studied, require that principals observe teachers three to four times as often as traditional systems and that they provide much more detailed, specific feedback to teachers. They are also expected to examine data such as student and parent surveys, assist teachers with self-reflection and goal setting, and be engaged with teachers' individual professional development plans. In addition, in the case of *Evaluate Davis*, principals are required to meet with each teacher individually three times yearly - at the beginning of the year to set goals, at mid-year to review progress, and at the end of the year, to complete a summative evaluation.

Unfortunately, the time allotted for principals to complete the more rigorous evaluation process has remained constant even though the workload has nearly tripled. This condition may increase the likelihood that principals will treat the process as routine, give quick, rather than meaningful, feedback, and rate all teachers as above average rather than risk engaging in time-consuming confrontations.

Quality teaching and student achievement. Debate continues among

researchers regarding what actually constitutes effective teaching. The absence of clear, definable targets for teachers to work toward obscures the formative goals of teacher evaluation. Darling-Hammond et al. (2012) noted little agreement among researchers and practitioners on what constitutes quality teaching even though there is consensus that teacher quality impacts student learning and that teachers' impact on student learning needs to be part of the teacher evaluation process. Moreover, teachers and administrators often fail to see the connections between evaluation, improved teaching, and student achievement gains, which also negatively affects their perceptions of evaluation systems (Schmoker, 2012). Goe (2013) and Towe (2012) each discussed the reality that insufficient research actually attempts to link current teacher evaluation practices to improved teaching and increased student learning.

Evaluate Davis

Evaluate Davis is typical of the new, more rigorous evaluation systems researchers and policy makers in general are looking to, and that are being implemented in most states in the U.S. today (Donaldson, 2016). It is based on this principle:

An effective evaluation system recognizes areas of strength and identifies areas for improvement by using multiple measures to assess performance based on standards and performance indicators. The process provides opportunities for continuous professional growth, feedback, and performance assistance, to ensure high quality staff and *learning first* for each student. (Davis School District, 2014, p. 1)

Evaluate Davis is an evaluation system with distinct differences from the system it replaced. The new process necessitates approximately three times as much time for administrators as the previous evaluation system, but also utilizes the recommendations

and requirements of policy makers and other stakeholders that educators are told have the promise of improving teaching and increasing learning, if the evaluation is implemented with fidelity (Goe et al., 2008). Table 1 displays the differences between the old and new evaluation processes.

Evaluate Davis requires teachers to engage with six components: self-assessment, observations, a professional growth plan, student growth measures, parent and student

Table 1

Differences Between the Previous System and Evaluate Davis

Evaluate Davis	Previous evaluation
Three yearly observations for career teachers and four for provisional teachers.	One yearly observation for career teachers and two for provisional teachers.
Yearly parent and student surveys for all teachers.	Parent or student surveys every fourth year for career teachers. Parent or student surveys yearly for provisional teachers.
Provisional teachers receive a summative evaluation twice per year. Career teachers have a full summative evaluation yearly.	Provisional teachers have a full summative evaluation yearly. Career teachers have a full summative evaluation every fourth year.
Four-point rating scale for observations: 1=Not evident; 2=Somewhat evident; 3=Evident; 4=Extraordinary. A 1 indicates the item was not seen during the observation; a 4 is rarely seen.	Three-point scale: Satisfactory, Needs improvement, or Unsatisfactory
Observations are electronically recorded so the teacher has immediate access to the observation. Each item has a place for comments for improved feedback.	Paper report with no room for comments.
The summative evaluation uses a four-point scale: Highly effective, effective, moderately effective, and ineffective.	Three-point scale: Satisfactory, Needs improvement, or Unsatisfactory
Much expanded and improved resources for professional development that are tied to the evaluation components.	No professional development

Note. Adapted with permission from “Educator Assessment System (EAS): Acknowledging and Honoring Quality Performance” revised June, 2009 and “Evaluate Davis: Licensed Educator Manual” updated 6/25/2014. Copyright 2014 by Davis School District.

surveys, and a summative evaluation. The process begins new each year but is intended to be ongoing throughout the year following a cycle that begins with a review of data that includes administrators' previous observations, parent and student surveys, and student achievement data. The data review is followed by reflection and self-assessment (teachers are expected to complete an electronic self-assessment which becomes part of the year-end summative evaluation), which leads to goal setting. (Teachers are required to complete a goal setting form that includes performance goals and student achievement goals.) Teachers then create a professional development plan based on their goals. (*Evaluate Davis* includes an extensive professional development resource called *EdPlus* that is aligned to specific standards and indicators in each domain that teachers are evaluated in.) An administrator reviews the self-assessment, goals, and professional development plan with the teacher early in the school year to assure alignment and provide assistance and encouragement. They will meet again mid-year to check progress toward goals and to review data from new observations and surveys. A summative evaluation is completed mid-year for provisional teachers. In a final meeting at the conclusion of the school year, data will be reviewed again and a summative evaluation completed (Davis School District, 2014).

Teachers are assessed on five domains: plan, teach, check, environment, and professionalism. Each domain is comprised of standards, which in turn include specific performance indicators. Observations, surveys, and self-assessments are aligned with the standards and indicators in each domain. Provisional teachers are observed four times each year and career teachers are observed three times. Observations are recorded

electronically and include comments and a checklist rating system based on a four-point scale: 1 = not observed, 2 = somewhat evident, 3 = evident, and 4 = extraordinary.

Observations are informal, unscheduled, and last at least twenty minutes. Principals are encouraged, but not required, to meet with teachers after each observation (Davis School District, 2014).

Research Questions

I completed this study seeking to gain an in-depth understanding (Creswell, 2013) of the attitudes, feelings, and dispositions that teachers and administrators approach *Evaluate Davis* with and how those attitudes and beliefs interact to impact implementation and outcomes. The research questions guiding my research were as follows.

1. What are administrators' and teachers' perceptions of the role *Evaluate Davis* plays in improving student learning gains?
2. What are the perceptions of teachers and administrators respecting how *Evaluate Davis* is used to improve teaching and learning?
3. To what extent do teacher and administrator attitudes regarding the efficacy of *Evaluate Davis* influence its perceived usefulness in improving teaching and learning?

Methodology

I used qualitative research methodology to evaluate educators' perceptions of the efficacy of a teacher evaluation tool, case study design (Stake, 1995) and multiple data-collection methods (Shenton, 2004), including focus group interviews and document reviews. The focus group interviews and document review allowed me to gather, analyze, and interpret in-depth qualitative data creating a descriptive picture (Creswell, 2013) of

the attitudes and beliefs of educators of a standards-based teacher evaluation system in the real-life setting of the Davis School District, a large suburban school district in Utah. The findings may provide those charged with using and implementing teacher evaluation systems with strategies and tools to aid the evaluation process and lead to better teaching and increased learning.

Researcher Bias

I approached this study from the perspective of a 37-year career educator in public schools. As a teacher for 10 years, I was observed and evaluated by the administrators I worked with. As an administrator for 27 years, I have observed and evaluated hundreds of teachers. I have also been involved directly in the formation and implementation of the evaluation system this study utilized to examine teachers' and administrators' perceptions and attitudes about evaluation. Although I endeavored to remove any bias I have from my interpretation of the data I collected, as the interpreter, the data was seen through my eyes as a participant-researcher. Furthermore, I have worked for the same school district for 36 years and was not unfamiliar to the administrators I interviewed. Likewise, some of the teachers who participated were familiar with me, although they had no employee-supervisor connection with me. However, my relationships with many of the participants may have facilitated trust and openness (Lincoln & Guba, 1985) that encouraged greater discussion and thus, more data. Frank discussions around the interview questions abounded and opinions were openly shared, not in spite of, but because of, the quality relationship between researcher and participant (J. C. Maxwell, 2007).

Significance of the Study

I engaged in this investigation with the goal of acquiring a deeper understanding of the attitudes, feelings, and dispositions that teachers and administrators approach *Evaluate Davis* with and how those attitudes and beliefs interact to impact implementation and outcomes of the evaluation process. An understanding of how educators' perceptions impact the evaluation tool's capacity to distinguish levels of teacher effectiveness accurately as well as to serve as an instrument for professional growth may influence the processes and strategies used in implementation. Many researchers believe there is great potential for teacher evaluation to have a positive impact on teaching and learning. However, that potential may be undermined if educators' attitudes and beliefs about the system are negative. The strategies and processes states and districts use in implementation are critical to assure that teachers and administrators see the potential and buy in to the process.

Through the examination of the attitudes, beliefs, and perceptions of educators toward the evaluation process, strategies for implementation may be clarified that may help assure the aims of the evaluation process are realized. An analysis of the current format for teacher evaluation implementation that acknowledges the influence of teachers' and principals' attitudes, beliefs and perceptions with the ultimate goal of teacher growth and increased student learning, may provide insight into procedures and policies that will further enable the aims of the system. For example, the methods used to train and prepare teachers and administrators for the evaluation process may need to be refocused and the methods used by districts to plan and implement professional

development opportunities may need to be restructured. Additionally, the dual roles of evaluation may need greater clarification for both administrators and teachers, in the context of the greater accountability movement. Furthermore, the results of my study may influence related research concerning evaluation and supervision and the extent to which comparable evaluation systems serve their intended purpose.

Summary

School administrators in Utah are required by state code to evaluate their teachers yearly using newly accepted, more rigorous standards that policy makers expect will elevate the quality of teaching and increase student learning. Complying with the new evaluation standards and requirements takes up to three times as much of administrators' time as previous evaluation models, yet there is little evidence that evaluating teachers actually improves the quality of teaching or increases student learning. In this case study, I examined the attitudes of teachers and administrators, in Davis School District, a large suburban district, towards the yearly teacher evaluation process, *Evaluate Davis*, in terms of its effect on student learning. Data gathered from group interviews and district-generated evaluation reports was analyzed to answer the following questions: What are administrators' and teachers' perceptions of the role *Evaluate Davis* plays in improving student learning gains? What are the perceptions of teachers and administrators respecting how *Evaluate Davis* is used to improve teaching and learning? To what extent do teacher and administrator attitudes regarding the efficacy of *Evaluate Davis* influence its perceived usefulness in improving teaching and learning?

The ensuing analysis adds insight into how teacher and administrator attitudes impact the effectiveness of the evaluation process in terms of student growth and school improvement and provides a platform for discovering practical strategies that may enhance the efficacy of the teacher evaluation process and make it of greater value to students, teachers, and administrators. The literature review in Chapter II serves to provide a theoretical framework guiding the study, historical background, and research basis for the study's main premise, that an understanding of how educators' perceptions impact the evaluation tool's capacity to distinguish levels of teacher effectiveness accurately as well as to serve as an instrument for professional growth may influence the processes and strategies used in implementation.

CHAPTER II

LITERATURE REVIEW

I designed this study to facilitate my examination of teachers' and administrators' current attitudes, beliefs, and perceptions regarding a standards-based teacher evaluation system, *Evaluate Davis*, recently implemented in the Davis School District, a large suburban Utah school district. I sought to understand the influence principals' and teachers' attitudes, beliefs, and perceptions have on the efficacy of the evaluation system in terms of instructional improvement, student learning, and school improvement. A primary goal was to illuminate approaches for implementation of standards-based teacher evaluation systems so that they may have their expected effect: better teaching, increased student learning, and improved schools.

The literature review includes the theoretical framework within which the study was conducted; the historical context in which the current evaluation system, *Evaluate Davis*, was studied, including the iterative improvements to the evaluation and supervision of teachers over the long history of public education in the U.S.; and a detailed description of the characteristics of past and contemporary evaluation systems, including the role of the principal, current issues, such as the often-conflicting summative and formative roles of evaluation, the traditionally poor perceptions of teacher evaluation held by many teachers and administrators, and the difficulty in defining quality teaching and connecting it to student learning. Finally, the literature review discusses current recommendations and expectations for teacher evaluation to impact teacher quality and student learning.

Theoretical Framework

Moore's *Theory of Transactional Distance* offers a conceptual strand to the general framework of this study. Although formulated to describe the psychological and communication distance (Moore, 1993) between teacher and learner that pertains to distance education programs, the construct of a transactional distance between teacher and learner applies to this research in the sense that it examines the perceptual distance between teachers and administrators and policy makers and researchers that is a potential concern in the implementation of teacher evaluation systems. Moore (1991) described transactional distance as a "space of potential misunderstanding" between teacher and learner. Giossos, Koutsouba, Lionarakis, and Skavantos (2009) defined transactional distance as a "distance of understanding between teacher and learner" that may lead to a "lack of mutual understanding or common perception of ideas, emotions...knowledge, approaches" and needs (p. 3).

Moore (1991) proposed that in "any educational program there is some transactional distance...in which the separation...is so significant that it affects behaviors in major ways" (p. 3). In other words, the degree of separation has a direct effect on how teaching is delivered. Moore describes "two sets of variables" or teaching procedures, dialogue and structure, which determine the magnitude of the transactional distance between teacher and learner. The kind of dialogue between teacher and learner is determined by the educational philosophy of the group who designed the course or program, as well as, by personalities and environmental factors. Structure, on the other hand, refers to elements in course design, the "rigidity of the programs educational

objectives, teaching strategies, and evaluation methods” (p. 4). It is described as the degree to which a program can be responsive to individual learners’ needs.

Moore’s *Theory of Distance* is applied to my research in very general, theoretical terms in the sense that teachers, as the recipients of the evaluation program, are comparable to the learners in distance education programs, and evaluators and the evaluation designers are comparable to the teachers and program designers in distance education programs. Considered in that sense, the concepts of “space of potential misunderstanding” and “lack of mutual understanding or common perception of ideas, emotions...knowledge, approaches” and needs, are clearly applicable. The idea that transactional distance between teachers and evaluators “affects behaviors in major ways” and how the degree of separation effects the delivery of the evaluation, is what I sought to learn through this research. The kind of dialogue between teachers and evaluators is evidently determined by the educational philosophy of the evaluation designers who created the evaluation system and is influenced by the personalities of teachers and evaluators as well as environmental factors in the individual work places. The structure, or rigidity of the evaluation tool’s objectives, implementation strategies, and methods clearly impact the degree to which the evaluation system can be responsive to individual teacher’s needs.

Rogers’ *Diffusion of Innovations* theory provided additional structural support to the theoretical framework of my research. Rogers (1983) defined Diffusion of Innovations as “the process by which an innovation is communicated through certain channels over time among the members of a social system” (p. 5). Valente and Rogers

(1995) described how sociological diffusion of new research paradigms follows a pattern in which new paradigms at first attract attention, but then are discarded for a newer paradigm. They clarified the stages in which new scientific paradigms are advanced within the social system. Initially, new paradigms attract a “furious amount of intellectual activity as scientists seek to test the new conceptualization” (p. 243). Over time, an “intellectual consensus” gradually develops among scientists and “interest eventually declines as fewer findings of an exciting nature are reported” (p. 243). During this stage the new paradigm continues to effect change as results of research are incorporated. Eventually, continued research exposes “anomalies in the existing paradigm” spurring another “scientific revolution (p. 244)” producing new paradigms.

Roger’s *Diffusion of Innovations* theory was foundational to my study on two fronts: First, the history of teacher evaluation follows the pattern described above by Valente and Rogers (1995). As the literature review demonstrates, evaluation systems have changed over time as educational paradigms have changed. And second, in this study I examined a teacher evaluation system, *Evaluate Davis*, which is a specimen of new evaluation systems being implemented throughout the U.S. The new systems are considered to be innovations in and of themselves, but are also based on the new paradigm of teacher growth as a means of improving student learning. As I sought answers to my research questions, I examined the perceptual complexities present in the widespread “diffusion” of a new innovation and paradigm within the education community (a new innovative method of evaluating teachers based on a paradigm of teacher improvement) and pursued additional methods of implementation and

sustainability (Valente & Rogers, 1995).

Greenhalgh, Robert, MacFarlane, Bate, and Kyriakidou (2004) defined diffusion of innovations as it pertains to health service organizations, but the definition is applicable to educational organizations also, as paraphrased: "...a novel set of behaviors, routines, and ways of working that are directed at improving *educational* outcomes, administrative efficiency...or users' experience and that are implemented by planned and coordinated actions." They explain that implementation refers to "...planned efforts to mainstream an innovation within an organization" and sustainability refers to "making an innovation routine until it reaches obsolescence" (p. 582).

Greenhalgh et al. (2004) described a continuum of approaches for diffusing innovations within health service organizations that is applicable to educational organizations such as schools and school districts. On the one end of the continuum is the "pure diffusion" approach in which the "spread of innovations is unplanned, informal, decentralized, and largely horizontal" or peer mediated. The outcome of this approach is unpredictable and emerges through the organizations social channels rather than from the top down. At the other end of the continuum is "active dissemination," an approach to spreading innovations that is "planned, formal, often centralized" (p. 601), and top down. The mechanism for dissemination is largely technical and comes from the top of the organizational hierarchy.

As the theory of *Diffusion of Innovations* applies to the implementation of a new teacher evaluation system based on a new paradigm, the approach for spreading the innovation in a school district is largely active dissemination. Specifically, in the case of

Evaluate Davis, certain attributes were mandated by the state legislature to assure that the evaluation would be in compliance with the most recent recommendations of researchers and experts. Davis district-level administrators carefully planned and carried out the actual implementation of the system at the district level and have worked to assure that it is implemented with great fidelity (the same in each school) in every school in the district. However, the implementation at the school level is somewhat haphazard, as some principals have put a lot of time and energy into training their faculty and preparing them for the more rigorous evaluation system, while others have taken the pure diffusion approach with little planning or training.

The *Marzano Evaluation Model* is another important part of the theoretical framework that supported my study. Marzano's model is built on the foundation of iterative improvements to the evaluation and supervision of teachers over the history of public education in the U.S. and on a strong research base supporting the following three points: (a) Teachers who are more skilled and expert produce greater gains in student achievement; (b) The evaluation process has the potential to help teachers to incrementally improve in skill and expertise; and (c) The methods for improving teacher skill and expertise are clear (Marzano et al., 2011). It is established on the principle that "the purpose of supervision should be the enhancement of teachers' pedagogical skills, with the ultimate goal of enhancing student achievement" (Marzano et al., 2011; p. 2).

The Marzano model has four domains: Classroom Strategies and Behaviors, Preparing and Planning, Reflecting on Teaching, and Collegiality and Professionalism. Within the four domains are sixty elements, forty-one in Domain 1, eight in Domain 2,

five in Domain 3, and six in Domain 4. The imbalance in elements indicates the model's emphasis on teachers' behavior in the classroom. The elements in Domain 1 are research-based strategies that have been shown to increase student learning and are further organized into three segments or subcategories: (1) routines segments, (2) content segments, and (3) segments enacted on the spot. The weight put on what teachers do in the classroom is a reflection of historical models of teacher supervision such as the Scientific Method proposed by Frederick Taylor in the early 1900s, Hunter's seven-step lesson design of the 1980s, and the Clinical Supervision models of the 1960s and 1970s (Marzano et al., 2011).

Domain 2, Planning and Preparation, deals specifically with research-based planning and preparation activities that relate to the classroom strategies and behaviors teachers use to improve student achievement. Specific subcategories include (1) lessons and units, (2) use of materials and technology, and (3) special needs of students. Domain 3, Reflecting on Teaching speaks to the need for teachers to self-assess their practice and to set goals for improvement in their skillful use of strategies and planning and preparation. Reflection and self-assessment were recommended by Dewey (1916) in 1916 and in 1933 as important to teacher growth and school improvement and were also important characteristics of the supervision models of the 1980s. Domain 4, Collegiality and Professionalism, is based on the premise that working well with others, in a professional and collegial manner, enhances the teachers' work in the other domains. Subcategories of Domain 4 are: (1) promoting a positive environment, (2) promoting exchange of ideas, and (3) promoting district and school development (Marzano et al.,

2011).

Summarily, the Marzano model of teacher evaluation is based on the principle that “evaluation is to enhance teacher effectiveness.” In practice, the model relies on multiple forms of feedback as the foundation for encouraging teacher growth in specific teacher skills. The system uses multiple measures of teacher growth in a “rigorous and informative approach” (Marzano et al., 2011; p. 103) that is intended to motivate professional development and recognize the different stages of progression toward expertise.

Historical Context

A brief historical sketch of teacher supervision and evaluation in the U.S. is included in the literature review to give the reader a clear view of education in America, as well as to promote reflection, and put the current status of education into perspective (Kliebard, 2004). For example, the current accountability movement in education has brought about changes in teacher supervision and evaluation policies and procedures, as well as the contemporary focus of curriculum, how we measure student achievement, and even the mission of public schools (Haefele, 1993; Popham, 2013). Researchers, stakeholders, and policy makers are better able to understand the current state of education in the U.S. if they know its roots and its progression. Better understanding of American education’s roots enables wise planning and decision-making for the present and future (Stearns, 1993).

From Monitoring to Improvement

Procedures for evaluating teachers' effectiveness have changed through time and generally reflect the current definition of what effective teaching is (Tucker & Stronge, 2005). The earliest forms of evaluation were supervisory and consisted mostly of monitoring teachers; little was done to improve teaching. During colonial times, teachers were considered servants of the community, so local governments used clergy and supervisory committees to hire teachers and monitor instructional quality (Sullivan & Glanz, 2009). There was little pedagogical knowledge and what was considered quality teaching was very different from what we experience today. Until the mid-19th century, supervision of teachers consisted of maintaining, rather than improving, current teaching standards (Mielke, 2012).

As school systems became more complex during the industrial revolution in the 1800s, the need for teachers who were specialists in particular disciplines became more apparent. Still, the goal of evaluation was supervisory rather than to improve teaching (Towe, 2012). Teaching began to blossom into a profession in the mid to late 1800s, and normal schools, where teachers learn to teach, were created. At the same time, the question of what effective teaching was became more important, so the job of evaluating and supervising teachers became more than simply monitoring their behavior; teachers were expected to know how to teach and to improve their teaching skills (Marzano et al., 2011).

Competing Purposes of Education

Competing visions of the function of education began to emerge in the late 1800s

and early 1900s. John Dewey had a big impact on education during the early 20th century, maintaining that schools should be organized around teaching the ideals of citizenship and democracy. He advocated a student-centered education that promotes integrating disciplines, connections to the real world, and differentiating instruction based on students' needs (Dewey, 1916). Frederick Taylor, however, believed in the scientific view of management, based on a "factory workers" (Marzano et al., 2011, p. 14) model in which specific behaviors are measured as a means of improving performance. Measurement and the analysis of data were encouraged as a means of providing specific feedback for teachers. John Franklin Bobbitt's (1918) book, *The Curriculum*, was based on Frederick Taylor's scientific model of management and became the norm for curriculum design in the U.S. (Towe, 2012).

Bobbitt was an advocate for The Social Efficiency movement, which maintained that preparing students for work and citizenship responsibilities was the purpose of education. Because students were certain to perform differing roles as adults, it was inefficient to teach them all the same curriculum. The curriculum advocated by the social efficiency movement incorporated components of general education for all students and differentiated elements of specialized vocational pathways based on students' intellectual capabilities, career objectives, and interests. A hallmark of social efficiency was the use of IQ tests to sort out who was and wasn't capable; it made no sense to waste time trying to teach someone who didn't have the aptitude (Kliebard, 2004).

As early as 1929, William Wetzel advocated measuring student learning, in addition to observing specific teaching strategies, to assess teacher effectiveness. Courses

were to have explicit, measurable learning objectives and teachers needed to learn to create and use reliable measures of student learning. Wetzel's (1929) scientific interpretation of education was primarily concerned with students learning specific content that could be easily measured using tests; the best teachers, according to this theory, would have the most students who meet prescribed learning objectives, as measured by standardized tests.

Current evaluation philosophy has its foundation on the footings of both Dewey's beliefs and the scientific model. For example, Dewey (1916) advocated for teacher reflection as a tool to improve teaching, and that practice resurfaced during the 80s (Marchiando, 2013) and is an important part of the standards-based evaluation systems currently employed in many states and districts throughout the U.S. today (Mielke, 2012). The scientific model is a major component, as well. Danielson and McGreal (2000) recommend that "evaluation systems emphasize student outcomes" (p. 19); the Marzano Teacher Evaluation Model includes student growth measures to assess teacher effectiveness (Marzano & Toth 2013); and federal education legislation, such as Race to the Top, requires that measures of student outcomes be a significant part of teacher evaluation, as well (Danielson, 2011; Donaldson & Donaldson, 2012).

The aims of public education are prone to broad pendulum swings. In the last 200 years we have entertained humanist/liberal education, the developmental movement, social efficiency, and the life adjustment effort, just to name a few (Kliebard, 2004). In the last 35 years there have been numerous pressures to change public schools to better meet their perceived purpose. The following reports have influenced how teachers teach

and are aimed at improving schools by either making learning more accessible to all students or by improving students' scores on measures of achievement on standardized tests: *A Nation At Risk* from The National Education Commission on Time and Learning in 1983 (Gardner, 1983) and the follow-up report in 1994, *Prisoners of Time*; The Carnegie Corporation's two reports, *Turning Points: Preparing American Youth for the 21st Century*, published in 1989 and the follow-up report, *Turning Points 2000*, published in 2000 (Jackson & Davis, 2000); and *Breaking Ranks: Changing an American Institution*, published in 1996 (there are actually three follow-up reports) by The National Association of Secondary School Principals (Maeroff, 1996). All have recommendations for how schools and teachers can improve and have therefore impacted the focus of teacher evaluation (Mielke, 2012).

The educational pendulum continues today to swing between Dewey-type ideologies, centered on citizenship and democracy, and the data-based, scientific model of management. The definitions of what an effective teacher is continue to reflect the prevailing beliefs about the purpose of education at the time (Kliebard, 2004). How we define what an effective teacher is impacts how teachers are evaluated (Tucker & Stronge, 2005).

Increased Emphasis On Evaluation

The mid 1900s and post-World War II era saw a shift away from scientific approaches as the purpose of education became more concentrated on the broader Dewey-type goals like citizenship, initiative, shared decision making, and democracy. At the same time, the importance of classroom observation as an integral component of

evaluating teachers became a point of emphasis, with the aim of evaluation now being formative; effective teaching and teacher growth became the goal (Marzano et al., 2011).

Clinical Supervision models. The 60s and 70s ushered in what Marzano et al. (2011) refer to as the “the era of clinical supervision”. Clinical supervision is a formative evaluation model characterized by pre-arranged classroom observations, analysis, a post-observation conference with the teacher, and a written analysis that is placed in the teacher’s file. The primary theory behind clinical supervision is that quality feedback from an experienced professional will influence teachers to try new instructional strategies. Three critical assumptions undergird clinical supervision models: (a) teachers view the evaluative relationship with administrators as an opportunity to be continuously involved in improving their teaching practice; (b) administrators have the skill and expertise to provide quality feedback to teachers; and (c) administrators will find the time, amid a host of other duties, to analyze observation data and turn it into meaningful feedback (Havens & Roy, 2007).

A change in focus. The year 1965 was an important one for public education in the U.S. The federal government legislated the ESEA as part of then President Lyndon Johnson’s War on Poverty. This Act initiated the process of changing the focus and purpose of schools in the U.S. away from sorting students into classes of who should get a job and who should go onto higher education, to that of learning for all, irrespective of economic status, social status, or race (DuFour & Marzano, 2011). Currently, ESEA puts \$11 billion per year into economically disadvantaged schools and neighborhoods in the U.S., with the goal of assuring that all students have access to high quality teachers and

the opportunity of high levels of learning (Towe, 2013). The focus on learning for all has impacted teacher evaluation. With the change in purpose has come evaluations that go beyond those things that are easily measured, like content knowledge, years of experience, and educational attainment.

Federal legislation, such as, No Child Left Behind (NCLB), Race to the Top, and the ESEA Flexibility Act, for example, have strong learning-for-all components, and are having a significant influence on teacher evaluation (Hazi & Rucinski, 2009). On January 6, 2015, ESEA was reauthorized when the Every Student Succeeds Act (ESSA) was signed into law. The stated purpose of ESSA is "...to provide all children significant opportunity to receive a fair, equitable, and high-quality education, and to close educational achievement gaps." ESSA has given states increased autonomy in designing teacher evaluation systems; what impact that will have on teacher evaluation reforms and influence on teacher quality will likely become increasingly important (Kraft & Gilmour, 2016). To meet the requirements of federal legislation and receive the funding that comes with compliance, forty-six of the fifty states, including Utah, have restructured their teacher evaluation systems with the intent of improving teaching (Donaldson, 2016; Popham, 2013). Contemporary standards-based evaluations include the following components recommended by evaluation researchers (Danielson, 2011; Goe et al., 2008; Marzano, & Toth, 2013; Pallas, 2010): More frequent teacher observations; more specific, detailed, immediate feedback focused on student learning; parent and student surveys; measures of student growth; and, teacher reflection and self-evaluation.

Until the post-World War II era, teacher evaluation was used commonly for

summative reasons, like employment decisions, compensation, and quality assurance. The clinical supervision models of the 70s were intended to be formative, however, leaving a void in the summative needs of school leaders (Havens & Roy, 2007). The systematic, formulaic characteristics of clinical evaluations were eventually found to lend themselves to making decisions about employment and compensation (Popham, 1988) and so schools and districts began using them for both summative and formative purposes.

During the 80s, The National Commission on Excellence in Education (1983) published *A Nation at Risk: The Imperative for Educational Reform* which spurred the accountability movement to new heights and served to intensify efforts to improve public education. This brought renewed scrutiny of the systems states and districts used to evaluate teachers under the assumption that better methods of evaluation would affect the quality of teaching and improve our schools (Hazi & Rucinski, 2009).

At the same time, more technical or didactic models of teaching emerged from the effective teaching literature (Hunter, 1986) and with that came models of evaluation and supervision that were also increasingly technical. Hunter (1984) developed a formative, and rather formulaic, seven-step model for lesson design that was focused on the ideal of teachers being continuously engaged in their growth as educators. Teachers were told to implement specific strategies and programs that were expected to raise the level of learning and put America back in the education race. The summative and formative purposes of clinical models, and the basic elements of Hunter's lesson design—anticipatory set, learning objectives, direct instruction, modeling, checking for

understanding, guided practice, and independent practice—persist as foundational pieces of many teacher evaluation systems in the U.S. today (Marchiando, 2013).

Today, standards-based evaluations are founded on the recommendations of federal legislation such as NCLB and Race to the Top, and are designed to achieve both summative and formative purposes (Danielson, 1996; Marzano & Toth, 2013). However, as Marzano (2012) reported, although the “vast majority of teachers believe that teacher evaluation should be used for both” (p. 15) summative and formative reasons, they also believe the emphasis should be on teacher growth and improvement.

Effects of No Child Left Behind. *No Child Left Behind* has changed the landscape of public education in the United States, increasing schools’ and teachers’ accountability for student learning and putting amplified focus on standardized testing and other measures of student achievement. The heavy reliance on standardized tests mandated in NCLB and used to evaluate teachers and schools have redefined learning and changed the meaning of quality teaching (Walsh, 2004). Because of NCLB, many states and districts evaluate both schools and teachers based on their students’ overall achievement on standardized tests, so teachers focus their instruction in ways that result in good test scores. Hazi and Rucinski (2009) pointed out that NCLB has compounded teachers’ and administrators’ accountability for improved student learning, targeting evaluation to upgrade teacher quality and improve students’ test scores. Teacher evaluation, they acknowledged, was further impacted as states have sought to achieve the NCLB goal of assuring that every student, irrespective of socio-economic status or ethnicity, is taught by a highly qualified teacher (bachelor’s degree in the subject being

taught and state licensed and certified).

In summary, due to ever increasing accountability standards for the education community, the focus of modern teacher evaluation systems is now on the relationship between teacher performance and student achievement (Marshall, 2005). Teaching and its relationship to student achievement is complex and not easily defined or measured, however (Tucker & Stronge, 2005). So it follows that modern day attempts to evaluate teachers have become more sophisticated, robust, and time consuming, using systematic, specific, research-based frameworks that experts say promote teacher growth and increased student learning (Marzano et al., 2011).

Characteristics of Contemporary Evaluations

Current literature offers several examples of evaluation models that promise to improve teaching and accomplish summative purposes (Danielson & McGreal, 2000; Havens & Roy, 2007; Marshall, 2005; Marzano & Toth, 2013; Popham, 2013; Toch, 2008; Towe, 2012; Weems & Rogers, 2010). Common characteristics of such evaluation models include: Based on explicit standards; use multiple measures of teacher effectiveness; continuous vs. once a year; founded on a relationship of trust and respect between administrators and teachers; frequent unannounced observations; focus on learning as the outcome; and, employ multiple levels to differentiate performance.

Darling-Hammond et al. (2012) noted that the best teacher evaluation systems are focused on helping teachers improve but also support timely and efficient personnel decisions. They recommend systems that require multiple classroom observations and

timely, detailed feedback by well-trained evaluators, self-reflection and goal setting, and multiple sources of data.

The National Governors Association (NGA) recommended six targets for creating teacher evaluation instruments that improve student achievement (Hazi & Rucinski, 2009): Clearly define teacher quality; include student learning measures; create professional accountability; train evaluators; expand stake holder participation in the design of evaluation system; and, focus evaluation on teachers' professional growth.

Focused on Learning

The purpose and vision of evaluation systems needs to be clear to evaluators and teachers (Marshall, 2005). Donaldson (2009) recommended teacher evaluations that are built on the considerable knowledge we have of what quality teaching looks like and aimed at increasing learning outcomes. Darling-Hammond et al. (2012) noted that evaluation should promote teacher growth and improved classroom instruction focused on increased student learning. Marzano et al. (2011) reported that the effectiveness of teacher evaluation systems rests on teachers' acceptance of the process as formative and their willingness to accept less than perfect ratings as a means of identifying areas of needed improvement. Danielson (2011) reported that teacher evaluations should focus on professional growth, with the intent of improving instruction; they should be aimed at motivating teachers to engage in professional development, learn new strategies, and make changes.

Teacher evaluations should produce the data that drives staff development and focuses attention on developing teachers' instructional ability to improve each student's

achievement (Towe, 2012). Havens and Roy (2007) also maintained that effective teacher evaluations should focus on learning as the outcome. Goe et al. (2008) recommended that evaluators work closely with teachers to use feedback to create a personal professional development plan that aligns carefully with effective teaching practices. They maintain that the central focus of evaluation should be to gather information that will be helpful in designing methods of improving instruction that improves learning.

Multiple Measures

Researchers recommend that successful evaluations include multiple sources of information and feedback, such as, principal observation, peer mentoring, portfolios, self-assessment, indicators of student learning, and student surveys (Marzano & Toth, 2013; Popham, 2013; Weems & Rogers, 2010). Towe's (2012) report recommended evaluations that are implemented with fidelity, use multiple data sources, and address the professional development needs of teachers, in order to encourage teacher growth. Donaldson (2009), investigated the components and implementation of teacher evaluations believed to positively affect student learning, noting that the combined benefits of Value-Added Measures, observations, portfolios, and performance-based assessment are proposed to have the most promise for influencing student achievement.

Kane et al. (2011) advocated for evaluation systems that incorporate both measures of student achievement gains and observations completed by trained professionals. Darling-Hammond et al. (2012) recommended systems that assess teacher quality based on multiple sources of data, as well. Goe et al. (2008) maintained that there is scant empirical evidence that any single measures, when used in isolation, are

appropriate methods of evaluating teachers' influence in improving student learning. They believe, however, that using multiple measures, which combine observation of teachers' practice with measures of students' academic growth, is one way to develop a comprehensive, valid, and reliable picture of a teacher's effectiveness. Measures they recommended include: observation, principal evaluation, self-evaluation, student ratings, and value-added measures (VAMs).

Student surveys. Many districts are using student survey data as part of their teacher evaluation systems (LaFee, 2014). LaFee reported in an article in *School Administrator* that students' perceptions of teachers' effectiveness correlate well with other measures, such as observations by trained observers and VAMs. Olatoye and Aanu (2011) argued that students, as the intended recipients of teaching, are in the best position to rate teachers' effectiveness. The majority of teachers, they found, see value in students' ratings as part of formative evaluations and in enhancing efforts to improve teacher quality.

Even so, due to reliability problems, student ratings are insufficient, Olatoye and Aanu (2011) warned, unless combined with other sources to evaluate teachers' performance. LaFee (2014) also noted that the usefulness of student surveys depends on the quality of the questions, the consistency with which the survey is administered to students, and the goals for which they are given. He recommended that survey questions should be consistent with the goals and philosophy of the school district and that survey data should be analyzed using tools that can assure accuracy. Additionally, he recommends that surveys be administered following consistent procedures with the goal

to improve instruction and student achievement.

Observations. Principals' observation of teaching has been a main ingredient of teacher evaluations since clinical supervision models were first introduced after World War II (Marzano et al., 2011). The rationale for observation is that principals' feedback will motivate teachers to make changes to their teaching. This rationale is based on assumptions that are seen by some as faulty, at best (Marshall, 2009). One assumption is that principals will spend adequate time in classrooms to make accurate judgments about the quality of teaching going on and will then have the expertise to give feedback to teachers that is specific and will motivate them to improve (DuFour & Mattos, 2013). Another is that teachers are looking to improve and will therefore accept principals' feedback as accurate and helpful (Havens & Roy, 2007; Steinberg & Sartain, 2015).

Weisberg et al. (2009) reported that principals do not spend enough time in classrooms to make accurate assessments of teachers' instruction. They noted that most observations are infrequent and completed with the intent of meeting the expectations of the evaluation process required by their district rather than to improve teaching. They noted a number of other faults with observation: (a) teachers expect to receive the highest possible ratings (and do more than ninety-four percent of the time); (b) the feedback teachers receive from their principals usually lacks specific suggestions for how to improve or what to do differently; (c) the professional development teachers receive is not aligned to data from the observations; (d) administrators are poorly trained observers and inter-rater reliability is suspect (In schools with multiple observers, observations will vary depending on who the observer is.); and (e) teachers lack trust in the observation

process and tend to feel picked on if they receive less than perfect ratings.

Dufour and Eaker (1998) and Marshall (2005) have argued that current evaluation procedures based on one to three observations per year have little influence on teachers' daily teaching practices and are usually not focused on student learning. Furthermore, they say, observations do not nurture an environment of candid, authentic, didactically productive dialogue between teachers and administrators. Their individual reports indicate little or no correlation between principals' observations alone and student learning gains. Steinberg and Sartain (2015) found support for concerns about teacher observation, noting that the evaluation system they studied in the Chicago Public Schools had only minimal effect in schools with higher performing, higher SES students, and zero effect in schools with lower performing students.

Measures of student growth. Darling-Hammand et al. (2012) reported increasing consensus among education practitioners, researchers, and stakeholders, that teacher evaluation systems should include measures of teachers' impact on student learning. Goe (2013) noted that most evaluation procedures implemented by districts and states in response to federal initiatives, such as Race to the Top, include measures of student growth as a main component. Marder (2012) noted that the federal government has been granting waivers from No Child Left Behind to states that include measures of teachers' impact on student learning. Marder also noted that VAMs, are now being utilized by many states across the country.

Darling-Hammand et al. (2012) have argued that VAMs “wrongly attribute to teachers” (p. 10) the impact of other influences on student performance, such as socio-

economic status, culture, and development. Their recommendation is that VAMs only be used for formative purposes and that they be only one of multiple components of teacher evaluation. Marder (2012) also expressed concerns that VAMs ignore the influence of information such as student mobility, that VAM data is very unreliable from year to year, and that several years of VAM data are needed to gain reliable results. Additionally, Value-Added Measures, according Marder, only measure one aspect of what schools hope to achieve. For example, school goals may include confident public speaking or raising graduation rates, which can't be measured by standardized tests. He also recommends that VAMs be used as one of many components of evaluation, with the recognition that the data provided is limited.

Feedback

Goe (2013) discussed the need for teacher evaluation systems that use quality data to generate specific feedback and continuous professional conversations, focused on growth, as the means for affecting student learning. Teachers, she found, have learned little from previous evaluation processes that provide limited, untimely, nonspecific feedback, and fail “to distinguish between effective and ineffective teaching” (p. 25). She further reported the connections between meaningful feedback and improved “planning, classroom management, formative assessment, differentiated instruction, and student-focused learning” (p.26) as well as increased self-reflection and ability to reach at-risk students. Marshall (2005) described authentic, data-based, face-to-face feedback and conversations, as critical characteristics of effective evaluations. Towe’s (2012) study found that both teachers and administrators agreed that feedback, collaboration, climate

of high expectations, respect, and rapport are key factors in successful evaluations. Darling-Hammond et al. (2012) also recommended timely, detailed feedback by well-trained evaluators as important to evaluations that motivate teachers to improve.

Specific standards. According to Goe (2013) meaningful feedback addresses specific teaching standards and expectations, learning standards where students are unsuccessful or are thriving, and daily teaching practice. Darling-Hammond et al. (2012) noted that professional standards such as those produced by The National Board for Professional Teaching Standards or the Interstate New Teacher Assessment and Support Consortium are the basis for evaluations that have proven to be much more reliable than statistical measures of student growth such as VAMs alone. To be useful, teacher evaluations must be based on more than standards, using checklists, observations, and ratings, according to Rennert-Ariev (2005). Effective evaluation, he notes, calls for teachers to engage in the examination of the usefulness of the standards or outcomes of their teaching and to reflect and collaborate about their practice, and claims that in practical assessment, teachers and administrators engage in dialog with the goal of refining understanding of teaching and learning.

Self-Reflection and Monitoring

Dewey (1916) recommended critical reflection as a component of teacher education that would enable teachers to improve schools. Dewey (1933) described self-reflection as “active, persistent, and careful” (p. 6) contemplation and critical evaluation of one’s self with a willingness to change. More recent research defines reflective teaching as characterized by a number of formal processes including keeping a daily

journal, peer observations and feedback, video recordings of lessons, or student feedback. Less formal methods may include simply having a conversation with a colleague, taking time to think about what went right or wrong with a lesson, or talking to one's self about concerns with a class (Cochran-Smith & Power, 2010; Mielke, 2012).

Rationale

Self-reflection has been suggested as a means of improving teaching ever since clinical models of supervision were introduced in the 70s (Hazi & Rucinski, 2009). Self-reflection is now widely endorsed as a component of standards-based teacher evaluation systems that focus on teacher growth and student learning (Danielson, 2011; Goe et al., 2008; Marzano & Toth, 2013). Hazi and Rucinski pointed out in their review of state's teacher evaluation policies, many states now require districts to embed teacher self-reflection in their evaluation systems and professional development practices.

Schon (1988) found that teachers improve in evaluation situations rich in reflective practice. Mielke (2012) recommended that reflection generates feedback that teachers may use to inform and improve practice. He explained further that a flaw with most current evaluation systems is that they fail to develop in teachers' dispositions of self-monitoring and self-reflection which lead to growth. Lakshmi's (2014) research concluded that teacher self-reflection has the potential to engage teachers in more meaningful professional development. Her study reported gains in teachers' personal and professional development through reflective practice and found that teachers became more aware of their teaching and more willing to improve their teaching practice. Danielson (2014) argued that teacher evaluation has the potential to have an influential

effect on teaching, but only “within the context of a collaborative observation/evaluation cycle in which the teacher plays an active role in self-assessment, reflection on practice, and professional conversation” (p. 32).

Role of the Principal

Principals are held accountable for how well they evaluate teachers and so expectations for principals’ leadership have been raised in recent years (Goe, 2013; Steinberg & Sartain, 2015). Steinberg and Sartain noted that the principal’s role has increased, from being primarily an evaluator to now including being an instructional coach, as well. To meet the demands of their new roles, Goe recommended principals need more training: (a) in gathering high-quality student-growth data based on an understanding of what good teaching is; (b) in analyzing and interpreting data, and turning it into recognizable, meaningful feedback that addresses specific teaching standards and expectations; (c) to provide quality feedback tied to specific standards with the goal of teacher growth; and (d) to become experts in maintaining positive relationships with teachers and in helping them monitor and evaluate their own growth and practices.

Marzano et al. (2011) pointed out that principals will need to take the time to be in classrooms much more and to provide frequent feedback in meaningful ways that teachers will acknowledge and that will assist them in planning to improve in needed areas. Towe’s (2012) research indicates that principals need to maintain positive relationships with teachers, and they need to help teachers see evaluation as useful and

fair and assure that evaluations take place in an environment of support, collaboration, reflection, and professional development. Halverson et al. (2004) suggested that evaluators and teachers need greater amounts of training on how to collect data, reflect on it, and present it in ways that increase its potential for improving teaching and learning. Additionally, they recommend training focused on helping evaluators develop their skill and will for giving critical feedback.

Current Issues

The literature describes numerous barriers to effective evaluation including, the sometimes conflicting roles of formative and summative evaluation; teachers' and administrators' negative attitudes and beliefs about evaluation systems; difficulty in clearly defining what good teaching is and then connecting the characteristics of good teaching to student achievement; and principals' lack of time, training, and will in completing meaningful evaluations. This section provides further detail regarding the barriers and explains the impact they may have on the evaluation process.

Dual Roles

The numerous reasons to evaluate teachers can be summarized to fit into two main categories: summative and formative (Popham, 1988). The current political climate of accountability demands that teachers be evaluated for summative reasons, such as recommendations for merit pay, rewards or advancement, to certify competency, to address weaknesses, and to continue employment (Pallas, 2010; Popham, 2013). The formative aims of teacher evaluation are equally important: First, to identify specific

practices that improve student learning and, second, to improve practice with the goal of improved student achievement (Darling-Hammond, 2012; Kane et al., 2011; Weems & Rogers, 2010). Havens and Roy (2007) described two models of evaluation: *Clinical supervision*, defined above as a formative model, and *performance evaluation*, which is summative and is less focused on continuous improvement. In performance evaluations, observations are unannounced, evaluators use a checklist of desired behaviors, and little or no follow-up is provided. Theoretically, teachers will respond positively to recommendations made during the evaluation.

Popham (1988) argued that summative and formative purposes of evaluation are not compatible; he claimed that summative uses undermine the formative. Gleave (1997) reported that attempting to use the same evaluation system to measure teachers' effectiveness and to help them improve only creates frustration and resentment. Goens and Lange (1983) advocated for the development of two evaluation systems, a summative system to assure quality and a formative system to enable teacher growth.

Darling-Hammond et al. (2012) noted that the best teacher evaluation systems are focused on helping teachers improve but also support timely and efficient personnel decisions. However, Popham (2013) described the impossible task of being both a summative and formative evaluator that school administrators are faced with. He provided evidence that when evaluation takes on a summative role, the formative role is weakened. He noted further, that when teachers feel threatened by summative consequences they are less likely to accept formative feedback in a positive way. Still, many experts, such as, Danielson and McGreal (2000), Darling-Hammond (2000), and

Marzano and Toth (2013) have influenced change in teacher evaluation systems in ways they hope will accomplish both summative and formative goals.

Poorly Perceived

Teachers and administrators don't always see the value in evaluation. However, their attitudes and beliefs about evaluation have an impact on the efficacy of the process. Marchiando (2013) found that the perceptions of teachers and principals toward the evaluation process have a direct impact on the system's implementation and effectiveness "as a tool for growth, which ultimately impacts student achievement" (p. 4).

Educators' attitudes. Teachers' and administrators' attitudes may be an important key to unlocking the potential for teacher evaluation to become an effective tool for improving teaching and increasing learning. Covey (2014) explained that how an individual or a group feels about something and their manner or disposition toward it would affect how they interact with it. For example, the following quote attributed to Henry Ford (*Readers Digest*, 1947) illustrates the idea: "Whether you think you can, or you think you can't, you're right" (p. 64). A goal of this study is to gain an in-depth understanding of Davis School District educators' attitudes towards *Evaluate Davis* with respect to the expectation that the evaluation system will have a positive impact on teaching and learning.

Nationally, teachers' views of evaluation have been somewhat negative (Pallas, 2010; Peterson, 2004; Popham, 2013). Halverson et al. (2004) found that both teachers and administrators viewed traditional evaluations as a perfunctory burden, used mostly to get rid of teachers with the poorest performance. They assert that the perceptions held by

teachers and administrators of the evaluation system determines how successful the evaluation procedures will be in affecting improvement in teaching and learning.

According to Darling-Hammond et al. (2012), there is general agreement among researchers and educators “that most current teacher evaluation systems do little to help teachers improve or to support personnel decision-making” (p. 8). They claim that traditional evaluation systems don’t do a good job of distinguishing between effective and ineffective teachers and they ignore important qualities of teaching that test scores can’t measure. In 2009, Hazi and Rucinski reported that traditional evaluation systems are faulty and difficult to use to make teachers more accountable, often exacerbating poor relationships between administrators and teachers. They discussed a lack of research connecting evaluation to improved teaching and increased student achievement and explain that is so partly because of the difficulty in gaining agreement among researchers and practitioners about what student achievement is. They also recommended the need for a cultural shift away from using evaluation as an accountability tool and directing the culture toward evaluation as a formative tool. They suggested that tying teachers’ salary increases to student growth data will complicate evaluations and threaten administrator-teacher relations. Furthermore, they claimed that state mandates regarding evaluation policies and procedures will reinforce the reputation that teacher evaluation is a meaningless ritual.

Like Hazi and Rucinski (2009), Halverson et al. (2004) reported a need for a cultural shift in schools and districts, from viewing evaluation as basically summative to a view that it is primarily formative. Teachers, they noted, were largely positive about the

feedback they received during their evaluations, and appreciated it more when it was given with the intent of improving teaching rather than for summative purposes. However, Havens and Roy (2007) found that although formative evaluations may benefit new teachers, who are receptive to feedback, veteran teachers are more prone to dismissing feedback and doing what they have always done, if consequences aren't attached.

Halverson et al. (2004) completed a case study in which teachers and principals were interviewed to identify their perceptions and attitudes toward the evaluation process. A range of perceptions came out of the interviews, from evaluation is “an opportunity to develop morale” to evaluation is nothing more than a “mandate” that creates a time-management problem for principals. They found that both teachers and administrators viewed traditional evaluations as a perfunctory burden, used in most instances to get rid of teachers with the poorest performance. The evaluation process, they found, did little to encourage teachers to improve and few teachers claimed to have made any changes in instructional practice as a result of evaluation. These poor perceptions of evaluation have had a negative effect on efforts to use evaluation to impact student learning in positive ways. Their study reported widespread belief that evaluation is “not a primary force for improving teaching (p. 12).”

The cycle. Pallas (2010) noted that current teacher evaluation systems have received criticism that they are “perfunctory and haphazard, relying on limited information” (p. 68) and that nearly all teachers receive top ratings. The perceptions of the evaluation system held by teachers and administrators determines how successful

evaluation procedures will be in affecting improvement in teaching and learning (Halverson et al., 2004). For example, Goe et al. (2008) reported that observation is a supportable evaluation measure to the extent that teachers and administrators believe the ratings actually indicate who is a good teacher. However, the perceived impotence of teacher evaluation procedures has led administrators to spend less time in classrooms and give little effort in providing constructive, formative feedback to teachers. Subsequently, valuable feedback to teachers has been lacking and teachers have found little formative worth (Zatynski, 2012) in the process, viewing evaluations as ineffective, as little more than an end-of-year hoop to jump through, or as tools to get rid of poor teachers (Peterson, 2004; Popham, 1998; Tucker & Stronge, 2005). Figure 1 illustrates how educators' perceptions impact the influence of evaluation.

Checklists. In the not so distant past, most teacher evaluations relied on checklists that were perceived as oppressive and did little toward teacher improvement or to

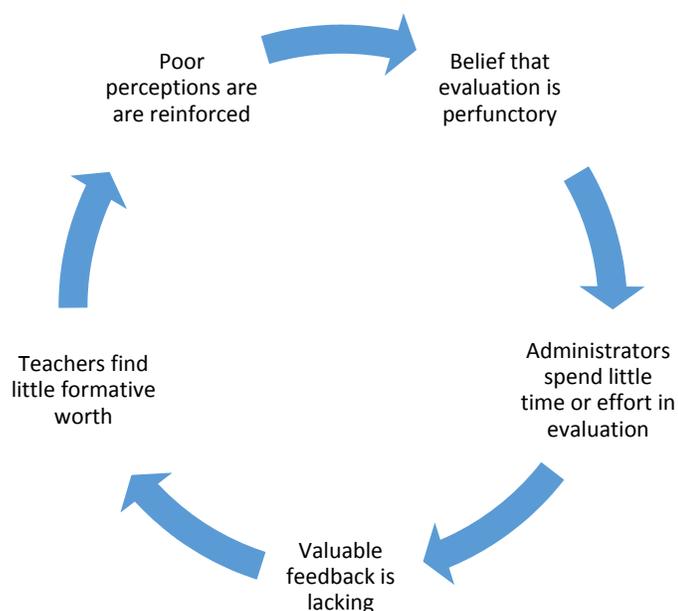


Figure 1. Evaluation perception cycle.

distinguish quality teachers from poor teachers (Danielson & McGreal, 2000; Weisberg et al., 2009). In 2008, Duffett, Farkas, Rotherham, and Silva reported that three out of four teachers felt the evaluation process in their school was merely a formality and had little impact on their teaching. This, they found, is particularly the case with checklist observations and when little feedback is given to teachers. Weiss and Weiss (1998) found that teachers and principals were increasingly frustrated with evaluation practices that stressed accountability for demonstrating specific teaching behaviors, like those of the Madeline Hunter model, instead of encouraging teacher involvement in self-growth and improvement of the school culture. They noted that many referred to evaluation as a meaningless exercise.

In their study of the Chicago Public Schools' teacher evaluation system, Steinberg and Sartain (2015, p. 72) reported that the 19-point observation checklist of classroom practices employed was "unpopular among both teachers and principals." The highest performing teachers felt that the evaluation did not provide meaningful feedback to improve instruction and most principals believed it didn't allow them to "adequately address teacher underperformance." Furthermore, the report noted no correspondence between teachers' ratings on the checklist and their final evaluations and that the final evaluations failed to distinguish between the ability of teachers to influence student learning.

To assure that educators take the evaluation process seriously, Goe et al. (2008) offered the following suggestions: Employ an evaluation that is both formative and summative; involve teachers in the process; assure teachers are aware of the targets they

are being evaluated on; provide teachers with quality feedback and an opportunity to discuss their evaluation; offer suggestions and support for areas targeted as needing improvement; and, ensure the process has been well explained and that teachers see the alignment of the evaluation with individual professional development, school and district goals, and school improvement plans.

The Principal Issue

Steinberg and Sartain (2015) noted the burden placed on principals to evaluate teachers' performance. They discussed the reality that the usefulness of the evaluation process rests with principals' skill in providing instructional guidance and their willingness to engage in the time-intensive process that includes large amounts of training and three to four times as many observations and post-observation conferences.

Halverson et al. (2004) reported that because principals carry the bulk of the burden of the evaluation process and play a crucial role in how components of the evaluation are implemented, they might also be partially to blame for teachers' poor perception of evaluations. Their attitudes are reflected in the attitudes of the teachers they supervise (Towe, 2012) as well as their willingness to spend time providing feedback, receiving training, or observing teachers. Without a strong belief in the worth of the evaluation process, principals are unwilling to invest time, energy, or resources in it (Kimball & Milanowski, 2009).

Toch (2008) described the typical teacher evaluation in public education as consisting of a "single, fleeting classroom visit" by an untrained building administrator wielding "a checklist of classroom conditions and teacher behaviors" that rarely focus on

quality instruction (p. 32). He noted that principals rarely give teachers an unsatisfactory rating, rarely use evaluations to improve instruction, and often don't even meet with teachers to discuss the results of their evaluations. Steinberg and Sartain (2015) also found that the efficacy of evaluation to improve teaching depended on two factors: the principal's competence in providing directed instructional assistance, and the teacher's willingness to accept and respond to instructional feedback.

Lack of time. Principals invest large amounts of time evaluating teachers and time is a big issue for principals, especially in secondary schools. Principals report a lack of time to observe teachers, for conferencing with teachers, and to provide quality feedback (Towe, 2012). According to Halverson et al. (2004) principals put a lot of effort into implementing all of the components of the evaluation, but find it time-consuming. To save time, principals resort to checklist observations, provide little or no meaningful feedback, and tend to give every teacher the highest possible rating. Many teachers, as a result, felt too little time was spent on providing meaningful feedback. Teachers reported that, to save time, principals' feedback is often too general and lacks the specificity to be helpful in improving instruction.

Steinberg and Sartain (2015) studied and reported on the Chicago Public Schools teacher evaluation system in their article *Does Better Observation Make Better Teachers?* The system requires that principals observe teachers three to four times as often as the previous system. Furthermore, administrators are expected to provide much more detailed, specific feedback; examine data such as student and parent surveys; assist teachers with self-reflection and goal setting; and be engaged with teachers' individual

professional development plans. According to Steinberg and Sartain, the new evaluation systems are accompanied by expectations that teaching will improve and student achievement will increase. However, they said, many principals and teachers wonder if the outcome is worth the effort; if expectations for the system are being met; or if the formative goals of evaluation are being realized?

Lack of skill. Principals have traditionally been poorly trained in evaluating teachers. Donaldson (2010) explained that a lack of direction from districts, including clear guidelines and rubrics for principals, contributes to poor quality evaluations. Halverson et al. (2004) found that principals rarely get feedback from their supervisors about whether they are doing a good job as an evaluator, even though relevant experience and training are seen as important to evaluators' success at assessing teachers. This lack of training results in principals being less than skilled in their ability to implement components of the evaluation, especially the critical elements of gathering and interpreting data and following it up with quality feedback (Goe, 2013).

Kimball and Milanowski (2009) reported that evaluators' knowledge and experience affects their accuracy. They found that evaluators with a greater degree of skill are more accurate in their assessments of teachers' strengths and weaknesses. They also reported that among evaluators who lacked training and experience, there was a tendency toward giving relative ratings or rating teachers in comparison to others. For example, if most teachers in a group are low performers, a teacher with a moderate performance will likely be rated higher than if she were compared to a group of high performers.

Lack of will. Principals often lack the will or motivation to implement all of the components of the evaluation process, including giving less-than perfect ratings (Tucker & Stronge 2005; Peterson 2004). Kimball and Milanowski (2009) found that principals, whose goal is to maintain good relationships, may be more lenient in their scoring, making the evaluation less effective in discriminating between teachers and identifying individual teachers' strengths and weaknesses. Donaldson (2010) referred to a culture in which principals would rather be seen as nice, than accurate, so they tend to rate all teachers as above average.

Donaldson (2009) explained that teacher evaluations have typically been ineffective in improving student learning outcomes, at least partially because nearly all teachers receive a satisfactory or better rating. A comprehensive study of teacher evaluation by The New Teacher Project (Weisberg et al., 2009) titled *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness* reported that in evaluation systems that assign either a satisfactory or unsatisfactory rating, 99% of teachers receive a satisfactory rating and in systems employing multiple levels of ratings, 94% of teachers received the top ratings.

Danielson (2001) noted that teachers and administrators have become accustomed to seeing the evaluation process as a validation and reward rather than a formative tool for improvement. Halverson et al. (2004) reported that principals rarely assign unsatisfactory ratings, giving mostly positive feedback, with only gentle criticisms. Although some (Marchiando, 2013) may attribute principals' failure to provide high quality feedback to lack of will or skill, a lack of clear rubrics that evaluators can refer to

quickly, without expending copious amounts of time, may be a big contributing factor, as well (Donaldson, 2010). Kimball and Milanowski (2009) concluded that the variations in principals' motivation, knowledge of instruction and the evaluation system, skill, and the school context influenced how they viewed and carried out their evaluation duties.

Connecting Teaching to Achievement

Educators' traditional poor perception of the evaluation process is likely related, at least in part, to the challenge of connecting student learning to the quality of teaching. Norman (2010) discussed the national effort to define clearly what comprises "accomplished teaching" (p.203) and what characteristics have a positive influence on student achievement. According to Norman, most assessments of teacher quality have concentrated on general pedagogical skills that hypothetically contribute to student learning but that haven't shown empirically that they truly do. Kimball and Milanowski (2009) also noted little empirical evidence that standards-based evaluations have an impact on teaching or learning.

In the article *Evaluating Teacher Evaluation*, Darling-Hammond et al. (2012) discussed the difficulty in connecting teacher quality to student achievement. They point out that teacher effectiveness is not a consistently defined concept, making it difficult to identify. Still, they note, the consensus among researchers and policy makers is that teachers' influence on student learning and the quality of teachers' practices need to be part of teacher evaluations. They reported that there is a problem with the assumptions that most evaluation systems make in attributing student-learning gains measured by test scores alone to teacher quality. There are a number of other important qualities of

effective teachers described in their report including: Deep understanding of subject matter; ability to connect learning to students' prior knowledge; able to create scaffolds and supports for learning; skilled at helping students make connections, apply, and practice new knowledge, and monitor their own learning; continually assess student learning and adapt teaching to students' needs; provide clear standards, constant feedback, and opportunities for revising work; and, successfully manage a collaborative classroom in which all students are engaged.

Hazi and Rucinski (2009) analyzed the statutes and department of education regulations for all fifty states to examine changes since the No Child Left Behind Act of 2001. Their report focused on three questions: What is the policy activity for teacher evaluation in state statutes and department of education regulations? How might these changes in statutes and regulations affect the practice of teacher evaluation? What implications for instructional supervision are likely to result from these policy actions? The NCLB goal of a "highly qualified" teacher in every classroom caused states to change policies about teacher evaluation. A Nation at Risk and various other national initiatives, they found, influenced changes in teacher evaluation that targeted evaluation to upgrade teacher quality.

Hazi and Rucinski (2009) found that evaluation holds the potential to improve teaching and that the focus on teacher improvement as a goal of evaluation has increased as a result of NCLB. The NGA responded to NCLB by including "defining teacher quality" as one of six targets for improving student achievement. As states have embraced NGA approaches for describing teacher quality, and embedded into state

statute and policy regulations evaluation practices that require professional development, the implicit principle of action is that improving teachers' instructional behaviors through professional development activities will lead to improved student achievement Hazi and Rucinski reported.

Quality teaching. School districts are faced with a daunting task when designing a teacher evaluation process. A general lack of consensus about what effective teaching is and what is important to measure makes evaluating it difficult (Goe et al., 2008). Darling-Hammond et al. (2012) explained that “teacher effectiveness is not a stable enough construct to be uniquely identified even under ideal conditions” (p. 9). Questions abound in the literature about how to define effective or quality teaching. Should teacher effectiveness be defined in terms of qualities such as those noted in No Child Left Behind (Weems & Rogers, 2010) like tenure, certification, or degrees earned? Should effectiveness be defined by how much impact (Marder, 2012) the teacher has on student growth as determined by value-added measures (VAMs)? How can evaluation systems accurately assess educators' qualities that are not easily measured or observed, such as their understanding of the subject matter or ability to connect new learning to students' prior knowledge?

Certification, experience, and knowledge. Some reports focus on commonly identified characteristics of teacher quality that align closely with the requirements of No Child Left Behind, such as, certification, subject matter and pedagogical knowledge, and teaching experience (Akiba et al., 2007). Goldhaber, Anthony, and Urban Institute (2005) investigated whether certification by the National Board for Professional Teaching

Standards (NBPTS) predicted teacher effectiveness with respect to elementary-level student achievement, and found that the “certification process itself” (p. 3) did nothing to increase teacher effectiveness. Walsh (2004) noted that “almost no empirical evidence shows that the [certification] process results in more effective teachers” and found that teacher certification is an “insufficient measure of a person’s qualifications to teach”. Walsh concluded that teachers may “prove their subject matter knowledge” to be deemed highly qualified under NCLB, but that doesn’t automatically translate into being highly effective (p. 23).

Value-added measures. A report from The Teaching Commission (2004) noted a growing trend to define teacher effectiveness narrowly by measuring student growth based solely on students’ scores on standardized tests. This trend is due, at least in part, to increased state and federal pressure for more accountability for teachers and schools for student growth. Many modern evaluation systems use sophisticated statistical models (Value-Added Measures or VAMs) to measure changes in students’ test scores over time to make a conclusion about the impact the individual teacher has on student achievement gains.

In their study of teacher evaluation, Darling-Hammond et al. (2012) claimed that Value-Added Models are based on faulty assumptions that are not supported by evidence and falsely attribute student gains on achievement tests to the teacher alone. They point to a number of other factors that influence learning, including: school, home and community, individual needs and abilities, health and attendance, peers, culture, prior learning and teachers, socio-economic status, and so forth. Furthermore, as noted above,

they identified numerous additional qualities of effective teachers that aren't measured by students' test scores.

Teacher qualities. Some define teacher effectiveness in terms of qualities like how much the teacher is able to close achievement gaps (Darling-Hammond, 2000). Goe et al. (2008) proposed a comprehensive, five-point definition of an effective teacher, intended to concentrate evaluation on multiple characteristics: Has high expectations of all students coupled with a willingness to help all students learn (self-efficacy); enables a positive academic and social learning environment which promotes regular student attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior; plans and structures engaging learning experiences, formatively monitors student progress, adapts instruction to student needs, and uses multiple sources to evaluate learning; values and promotes civic-mindedness and diversity; collaborates effectively with colleagues, parents, and administrators to ensure the success of all students.

Darling-Hammond et al. (2012) suggest qualities such as the teacher's ability to build successful scaffolds and supports for learning; engage students with metacognitive strategies; continuously assess and adapt to students' needs; provide specific standards, continuous feedback, and chances for students to redo their work; and promote an inclusive, collaborative classroom in which all students are engaged.

Norman (2010) addressed the relationship between assessment of quality teaching and student-learning and the challenges associated with making those connections, noting the efforts of educators and policy makers to clearly delineate what comprises quality

teaching and what teacher characteristics have a positive influence on student achievement. The result has been evaluation procedures aimed at the ability of teachers to impact student learning. Norman's report argues that assessing teacher quality must go beyond the "highly qualified" mandates of NCLB, and the general pedagogical skills believed to influence student learning, to include the teacher's impact on students' growth in terms of knowledge, skills, and understanding.

Ziebarth-Bovill, Kritzer, and Bovill (2012) reviewed literature which validates two points: there is no magic formula to identify great teachers and great teachers' impact student achievement. They proposed two general categories of characteristics important to being an effective teacher. First, is a professional skill-set that includes academic preparation, content knowledge, and pedagogical ability. The second deals with the social aspect of teaching: builds connections with kids; believes in his ability to get kids to learn; believes in all students' ability to learn; keen sense of awareness; high expectations of all kids; and empathy.

Teacher evaluation and teacher quality. A requirement of No Child Left Behind is to ensure teacher quality, with the expectation that student achievement will improve. Porter-Magee (2004) discussed the much-debated highly qualified provision of NCLB that recognizes the connection between teacher quality and student achievement and endeavors to assure all students have access to highly qualified teachers. By mandating minimum standards for teachers, Porter-Magee pointed out, NCLB opens the door to evaluating teachers by student outcomes rather than years of experience or degrees earned.

Weems and Rogers (2010) pointed to effective teacher evaluation as the means to promote teacher quality. In their definition, “quality encompasses such dimensions as intellectual rigor in courses, skill in elucidating difficult material, accessibility to students, interest in students’ academic progress, and the ability and willingness to assess that progress thoroughly and fairly” (p. 19). They argue that teacher quality is important to student learning gains, and teacher evaluation is the path to improving quality. Porter-Magee (2004) proposed a similar relationship, citing the work of Rivers and Sanders (2002) who reported the lasting impact of both good and poor teachers on students’ academic achievement.

Definition of achievement. Darling-Hammond et al. (2012) noted that in addition to the difficulty researchers have found in defining quality or effective teaching, they have also lacked consensus in defining student achievement. Goe et al. (2008) described the close alignment between teacher effectiveness and the concept of student learning, explaining that both are somewhat nebulous in terms of a standard definition. The U.S. Center for Public Education (2005) defined student learning narrowly as improvement on state-mandated tests. However, Gratz (2011) concluded, in addition to their easier-to-measure academic goals, schools have many other important aims: social and emotional growth, becoming good citizens, appreciation for the arts, teamwork and collaboration, critical thinking, problem solving, communication skills, and a good work ethic. Such outcomes are more difficult to measure and attribute to teacher effectiveness.

Expectations

Because of accountability pressures to ensure learning for all students,

stakeholders need to know if formative evaluation processes impact teachers' growth, and as a result, student learning (Marshall, 2005). Greene (1992) discussed "the notion that supervising and evaluating teachers might lead to their professional development" (p. 131). Do current teacher evaluation systems identify and measure, as Towe (2012) posited, "...the instructional strategies, professional behavior, and delivery of content knowledge that affect student learning" (p. 8)? On the one hand, a 2009 report titled the "The Widget Effect" (Weisberg et al., 2009) concluded that teacher evaluation processes in the U.S., in which teachers are given the highest possible ratings by their evaluators 94-98% of the time, fail to differentiate between poor, average, good, and excellent teaching, treating all teachers as virtually the same. On the other hand, there is the expectation of policy makers that teacher evaluation is the key to improving our nation's schools (Darling-Hammond, 2012; Goe et al., 2008; Goldrick, 2002; Popham, 2013).

Improve teaching. Quality teaching matters to student learning, Goe (2013) noted, and one goal of evaluating teachers is to gain understanding of what effective teaching is and to develop approaches to improve instruction. The major force behind the current revamping of teacher evaluation systems across the United States is the belief that if we can improve teacher quality, we will raise student achievement (Goe, 2013; Popham, 2013). Akiba et al. (2007) noted "Teacher quality is widely recognized by policymakers, practitioners, and researchers alike to be the most powerful school-related influence on a child's academic performance" (p. 369).

A large body of recent research has come to the conclusion that the quality of instruction students receive is one of the most important variables in determining how

much they will learn; maybe even as important as home and socio-economic factors (Buddin & Zamarro, 2009; Danielson, 2001; Darling-Hammond, Wise, & Klein, 1999; Hattie, 2009; Haycock, 1998; Marzano, 2003; National Commission on Teaching and America's Future, 1996; Sanders & Rivers, 1996; Tucker & Stronge, 2005). Cummings (2002) noted, "The classroom teacher, more than any other single variable in the education equation, can enable all students to learn" (p. 4). Wang, Haertel, and Walberg (1993) reported that classroom teachers have greater effect on student learning than factors like state, district, and school policies. Mendro (1998) indicated that effective and ineffective teachers have an influence on their students, for good or bad, for several years to come.

Goldhaber (2002), a senior research associate at the Urban Institute in Washington, DC, examined the literature regarding quality teaching, seeking to identify the attributes of teachers who positively influence student-learning outcomes. He identified several studies that found that effective teachers impact student achievement more than other outside factors such as home environment and socioeconomic status. In a quantitative analysis of states' achievement data, Darling-Hammond et al. (1999) explored how teacher qualifications and other variables like class size are related to student achievement. She noted that quality teachers are more likely than class size, poverty, language background, or minority status to be the determining factor in student achievement.

Borman and Kimball (2005) concluded that teacher qualities are responsible for as much variance in student learning as any other school characteristic. Other reports,

(Akiba et al., 2007; Porter-Magee, 2004) identified characteristics of teacher quality that align closely with the requirements of No Child Left Behind—state certification, subject matter and pedagogical knowledge, and experience—and found varied correlation with increased student achievement.

With research showing that teachers are indeed the most important school-based factor in the learning process, improving the quality of teaching has become a focal point for policy makers. Darling-Hammond et al. (2012) noted that the consensus among researchers and policy makers is that teachers' influence on student learning and the quality of teachers' practices need to be part of teacher evaluations. For instance, Goldhaber (2002) maintained that if our goal is to improve our nation's public schools, then we need to improve the pool of teacher candidates so that principals can hire better teachers. As Whitaker (2004) noted, it's not what needs to improve, it's who. Likewise, DuFour and Marzano, (2011) stated, "At a very basic level, school reform is about substantively changing people..." (p. 47).

Evaluation to the rescue. Standards-based teacher evaluation systems are perceived to be powerful tools and, teachers and administrators are told, hold great promise to improve teaching and learning in schools (Danielson & McGreal, 2000; Halverson et al., 2004; Odden & Kelley, 2002). Toch (2008) reported "through their focus on the quality of teaching, teacher evaluations are at the very center of the education enterprise and can be catalysts for teacher and school improvement" (p. 32). Hazi and Rucinski (2009) suggested that the theory that evaluation and supervision will improve teaching has gained a larger following recently. They noted the NGA

recommendation that requirements for professional development activities that increase teachers' professional behaviors be imbedded in states' statutes and policy regulations. Kimball and Milanowski (2009) acknowledged that the current political climate encourages the use of evaluation to make high-stakes decisions regarding teacher salary and retention. They posited that administrators' evaluations of teachers might actually have some validity, justifying their use for making employment and salary decisions.

Havens and Roy (2007) promoted the perspective that teachers' professional growth and improvement is influenced by the evaluation process, and the end result is increased student learning. Teacher evaluation, for many stakeholders and policymakers, is the key for improving teacher quality and hence, our public school system (Goe et al., 2008). In a recent article, *The Right Start: Creating a Strong Foundation for the Teaching Career*, Darling-Hammond et al. (2012) discussed the political forces—federal, state, and local policy makers—that are looking toward teacher evaluation impacting student achievement and as the way to improve public education. In response to political influences, experts such as Goe et al. (2008), Marzano and Toth (2013), and Danielson and McGreal (2000) have proposed research-based teacher evaluation processes they believe have promise for improving student learning.

The impact of stakeholders. Scrutiny and criticism of our nation's public schools has been increasing since *A Nation at Risk* was published in 1983. NCLB created increased attention from the public and school personnel, and most recently, two federal initiatives—Race to the Top and the ESEA Flexibility Program—have sparked radical changes to how public school systems do business, and specifically how they plan to

evaluate teachers (Popham, 2013).

Race to the Top was initiated in 2009, offering huge grants to states if they would incorporate teacher evaluation procedures that were more rigorous, using student achievement gains as a major factor. The ESEA Flexibility Program (U.S. Department of Education, 2012) came about in 2011 as a way for states to avoid NCLB sanctions; it too required states to give their teacher evaluation systems a facelift. Both programs have similar requirements of states' teacher evaluation systems to qualify: a focus on continual instructional improvement, student growth, regular, useful feedback, and multiple sources of evidence (Popham, 2013; Steinberg & Sartain, 2015).

The current emphasis on evaluation in the U.S. certainly leads to the conclusion that policy makers see it as a key to improving schools and teaching. Race to the Top, The ESEA Flexibility Act, and NCLB, indicate hopes of teacher evaluations consistent with those of the NGA; they expect evaluation to pave the way for achieving the goal of a "highly qualified" teacher in every classroom (Hazi & Rucinski, 2009). The NGA acknowledged six improvements to teacher evaluation systems they believe will result in greater student achievement: define teacher quality, make student learning measures part of teacher evaluation, create professional accountability through career ladders, train evaluators, and expand stake holder participation in the design of evaluation systems.

Summary

Teacher supervision and procedures for evaluating teachers' effectiveness have changed through time and generally reflect the current definition of what effective

teaching is. From Colonial times to the present, as teaching in the United States has grown into a profession, and consequently become more complex, the methods of evaluation have evolved to be more technical and didactic. The footings of current standards-based evaluation systems include philosophies and ideals from Dewey's era, such as critical reflection and the scientific model of management, as well as Madeline Hunter's seven-step model of lesson design.

Federal legislation has shaped the education landscape and effected changes in how and why teachers are evaluated. The ESEA of 1965 spurred a change in the focus of public education, away from sorting students into classes of who should get a job and who should go onto higher education, to that of learning for all. Consequently, interests into the effect individual teachers have on student learning began to grow. Clinical supervision models of evaluation, which focus on specific teacher behaviors believed to be associated with increased student learning, resulted.

The publishing of *A Nation At Risk* in 1983 set in motion a movement to not only make teachers accountable for how they teach, but also for student achievement results. Since then, policy makers and stake holders have looked increasingly to teacher evaluation as the way to improve student achievement in public schools. In the last two decades, federal legislation, including NCLB, Race to the Top, and the most recent Every Student Succeeds Act, has increased educators' accountability for student learning and forced dramatic changes to how teachers are evaluated.

Current standards or performance-based evaluation systems are now required in many states to meet expectations of policy makers. The common elements of standards-

based evaluation systems include: focus on instructional improvement and student learning growth; self-reflection and goal setting; frequent quality feedback; and multiple sources of evidence that includes observations of teaching, parent and student feedback, and measures of student achievement.

The history of teacher supervision and evaluation, and the rationales (formative or summative) for evaluating teachers, have an impact on the methods and systems states and school districts employ. Those who evaluate teachers have perceptions, beliefs, and attitudes about the process that influence how they perform their obligation. The teachers who are being evaluated also have perceptions, beliefs, and attitudes that influence their reception of the evaluation process and the accompanying feedback. The aim of this study is to illuminate educators' attitudes, perceptions, and beliefs and examine the influence they have on the evaluation process. By having a greater understanding of educators' attitudes, beliefs, and perceptions of evaluation, the methods for implementing evaluation systems can be altered to provide the best prospects for success.

CHAPTER III

METHODOLOGY

The qualitative research methodology and procedures (Creswell, 2013; Stake, 1995) I used to study how teachers' and administrators' perceptions, beliefs, and attitudes impact the implementation of the standards-based evaluation system, *Evaluate Davis*, are described and explained in the following sections: (a) design, (b) research questions, (c) data collection, (d) case selection, and (e) data analysis.

Design

I conducted this qualitative research using a case-study approach (Yin, 2003) as the primary method for gathering data. Data gathered from both administrator and teacher focus group interviews, as well as from a review of Davis School District teacher evaluation training documents and reports, were analyzed to investigate the attitudes, beliefs, and perceptions of teachers and administrators in Davis School District regarding the current teacher evaluation system, *Evaluate Davis*. Elementary and secondary teachers and administrators participated in the focus groups. A descriptive analysis of the qualitative data gathered in the case study permitted me to identify and clarify the attitudes, beliefs, and perceptions of both principals and teachers regarding the implementation of *Evaluate Davis*, as well as to compare the responses from both focus groups. Furthermore, the case study approach allowed me to deeply examine how teachers' and administrators' perceptions of the evaluation process influence implementation of *Evaluate Davis* (Yin, 2003).

The question of whether the evaluation system, *Evaluate Davis*, is perceived by teachers and administrators to be used effectively, and how their perceptions effect the implementation of the system, is the problem that I investigated for deeper understanding (Creswell, 2013). Davis School District, as the case chosen, provided a real-life, bounded (Merriam, 2009) environment in which I examined the impact of teachers' and administrators' perceptions on the implementation of a standards-based teacher evaluation system.

Research Questions

The research questions that guided my research were as follows.

1. What are administrators' and teachers' perceptions of the role *Evaluate Davis* plays in improving student learning gains?
2. What are the perceptions of teachers and administrators respecting how *Evaluate Davis* is used to improve teaching and learning?
3. To what extent do teacher and administrator attitudes regarding the efficacy of *Evaluate Davis* impact its usefulness in improving teaching and learning?

Data Collection

In this section, I describe the two primary data-collection methods I used, document review and focus group interviews, in accordance with accepted principles of qualitative scientific research in education (Creswell, 2013; Shavelson & Towne, 2002). Multiple data-collection methods, focus group interviews and document review, contributed to the trustworthiness of the study (Shenton, 2004)

Documents

The documents I reviewed included two types of district-generated documents: First, documents relating to training teachers and administrators regarding how to implement the evaluation tool with fidelity, such as, training documents and manuals for administrators and teachers, Power Point presentations used by administrators to train teachers about the evaluation process, leader notes that accompany the slides in the Power Point presentations, e-mails and other communications from district-level leaders. The school district also provides rubrics for each of the five standards measured in the evaluation: Plan, Teach, Check, Environment, and Professionalism. And second, teacher observation reports available to administrators in Davis School District that provide information regarding the fidelity with which administrators and teachers have implemented *Evaluate Davis* in the school. These documents provided me with information unavailable through the focus group interviews and added insight and perspective to the interview data (Bogden & Biklen, 2007). They provided context and spawned new questions, creating greater understanding of the influence of teachers' and administrators' perceptions of evaluation on the implementation of the process (Glesne, 2011). Below is a brief description of each of the documents I reviewed. Actual examples are included in the appendices. A permission letter from Davis School District to use *Evaluate Davis* documents may be found in Appendix A.

The *Evaluate Davis Observation Item Report* compares the ratings of the school to the ratings of the district, across the 34 indicators in the observation protocol. This report was fundamental to this study because one main concern with previous teacher

evaluations has been what some education researchers refer to as the “widget effect” (Weisberg et al., 2009), which indicates the tendency for administrators to give all teachers the same high ratings. The “widget effect” makes the evaluation process of little worth in terms of differentiating between poor, good, and excellent teachers, or for providing focused professional development.

Evaluate Davis observer item report. This report lists all of the school administrators in the district, by school, and shows the percent of 1s, 2s, 3s, and 4s each administrator gave, the percent total for the school, and how many observations each administrator completed and the total completed for the school. This report was of value because it provided information about the fidelity with which the evaluation instrument, *Evaluate Davis*, was administered in each school and by each administrator within each school.

Evaluate Davis observation completion report. This report lists each school in the district and shows the percent of teachers with at least one, two, or three observations, and the average observation rating. The *Observation Completion Report* also provides information regarding fidelity of implementation. Knowing the average observation rating allowed me, the researcher, to make important predictions about how educators may feel about the evaluation process.

Classroom observation detail report. The fourth document of interest is the *Classroom Observation Detail Report*, which lists each teacher in the school and gives the number of times they were observed along with the percent of ones, twos, threes, and fours given. Permission was needed from participating teachers for me to view this

report. Knowing if a teacher was rated high or low in observations was helpful as interview data was analyzed and I attempted to interpret and bring meaning to it.

Demographic information. In addition to the above documents, demographic data, gathered using the form displayed in Figure 2, helped inform the study regarding the age and tenure of participants.

Demographic information was important to this study because it showed possible relationships between factors such as age and tenure and the attitudes and perceptions that are illuminated through the group interviews (J. A. Maxwell, 2012). For example, many researchers have found that new teachers are more prone to accept and value feedback from administrators than experienced teachers (Popham, 2013).

Focus Group Interviews

Focus group interviews (Holstien & Gubrium, 2003) are recommended by Creswell (2013) when the interaction among participants will produce the best data and interviewees are comparable and cooperative. Glesne (2011) found that participants in focus groups are often engaged and eager to share multiple perspectives with a great degree of “depth and intensity” (p. 131). Barnett (2006) recommended interviews

Name _____
Age _____
Number of years teaching _____
Subjects taught _____

Figure 2. Demographic information form.

because they allow for probing and clarification of data while giving opportunity for in-depth responses. Furthermore, focus groups are very efficient in terms of the time (one hour for each group) and the amount and quality of data gathered (Morgan, 1997).

I presented open-ended questions in separate focus group interviews with teachers and administrators, and responses were digitally recorded, transcribed, and analyzed. Teachers were interviewed separately from administrators to assure that they were candid and felt comfortable sharing their perspectives among themselves and with the researcher (Patton, 2002). Although there may have been some question regarding the openness and frankness of the responses generated in the interviews with teachers and administrators, I had no supervisory authority over any of the participants, and they were assured that all responses would be kept in strict confidence and all respondents would be de-identified.

To provide consistency and assure validity (Creswell, 1998), a semistructured interview protocol (Glesne, 2011; Gubrium & Holstein, 2002), consisting of six open-ended questions, was used in the interview process (see Appendix B). Bloor and Wood (2006) explained that focus group interviews tend to engage participants in a form of debate facilitated by the questions posed by the researcher. Furthermore, qualitative studies are by definition exploratory (Bogden & Biklen, 2003) so, although the focus group interviews began with the same questions, the intent was to have them evolve into controlled discussions that varied between the teacher group and the administrator group.

The outcome of the interviews was to learn of participants' experiences, opinions, and thoughts concerning the evaluation process and what place, they believe, the process holds in their schools' over-all success in student learning. Questions, as seen in the

Interview Protocol in Appendix B, were open-ended and led to discussions about the various components of the system. There are six components of *Evaluate Davis*: (a) classroom observations and feedback from administrators; (b) parent and student surveys; (c) self-evaluation and reflection; (d) professional development; (e) goal setting; and (f) measurements of student learning. Initial questions were followed up with a variety of probing questions, dictated by the tone of the discussion and participants' comments, leading to an increasingly detailed perspective of how each component is administered and the attitudes and feelings of the participants regarding the components.

The setting of the interviews is important to the extent that participants need to feel comfortable and assured they can speak freely and openly and that they may choose not to speak if they prefer (Morgan, 1997). The teacher focus group interview was held in a conference room at a local public library after teacher contract hours. The conference room was of adequate size, with comfortable seating arrangements, in which all participants could see, hear, and understand each other. The administrator focus group interview was held in a conference room in a junior high school at a time convenient for all six administrators.

Participant Selection

This section includes a description of how participants were selected for focus group interviews. Focus group interviews are not intended to provide results that can be generalized to the entire population. Given (2008) suggested that the method of choosing samples in qualitative studies is dependent on the goals of the study and Yin (2003) noted that once researchers decide what they want to learn from their research, the appropriate

sampling strategy becomes apparent.

Purposive sampling. Purposive sampling is common in qualitative research and implies that the researcher intentionally chooses the sites and participants that will best meet the objectives of the study (Palys, 2008). In case studies, researchers are interested in attitudes and beliefs of individuals and groups, and the ways those attitudes and beliefs interact within the cases being studied (Merriam, 1988). The teachers and administrators invited to participate in this study were intentionally selected with the goal of gaining a better understanding of the social and psychological aspects of the teacher evaluation process and how the attitudes of both the evaluator and the person being evaluated impact the process in ways that may improve teaching and learning.

Maximum variation sampling. There are numerous purposive sampling strategies described in the qualitative research literature (Becker, 1998; Stake, 2005). Maximum variation sampling is the purposive sampling strategy that most closely aligns with the goals of this study. While seeking to describe the broad range of attitudes and beliefs that impact the evaluation process, I looked for participants whose views included the whole range of perspectives, from the extremes at both ends of the spectrum, plus the variations between the extremes, as explained by Palys's (2008) description of maximum variation sampling.

Following maximum variation sampling strategies (Creswell, 2008), participants were intentionally selected whose attitudes likely reflect the spectrum of attitudes and beliefs, as described by Palys (2008), that would be normal within the general population, and who would likely provide me with the best information regarding perceptions,

attitudes, and beliefs about teacher evaluation. Those individuals were recommended to me through discussions with the district's Research and Assessment director, elementary and secondary School directors, and school principals. Upon receiving recommendations, I used email to invite each individual to participate in a 45-minute focus group interview. I followed up invitations with an individual face-to-face interview to assure each understood the purpose of the study and what participants would be asked to do. For all participants, each expressed a desire to share perceptions of the evaluation process, both positive and negative. Tables 3 and 4, discussed and shown in Chapter IV, provide demographic information for each participant.

Confidentiality. Research records were kept confidential, consistent with federal and state regulations. Only I had access to the data that was kept in a locked file cabinet, on a password protected computer, in a locked room. To protect privacy, personal, identifiable information was removed from study documents and replaced with a study identifier. Identifying information was stored separately from data and was not used in the report. The study was completed in May, 2017 and all personal identifiable information was destroyed at that time. Audio recordings were also destroyed at the completion of the study. Participants were asked to sign the Informed Consent Form (included in Appendix C) upon agreement to take part in this study.

Disclaimer. Purposive sampling includes intentionally selecting participants (Creswell, 2008) that assist the researcher in studying a central issue or phenomena. The task of intentionally selecting participants who fit the description my research was targeting was facilitated by the quality relationships I have established with the district-

level administrators and principals who lead the district. Lincoln and Guba (1985) described how sustained relationships between researcher and participants enable an ample understanding of the organization being investigated and create trust between the parties. My relationship with my principal colleagues served to create an environment in which they were open and forthright in their comments during the focus group interviews. Frank discussions around the interview questions abounded and opinions were openly shared, not in spite of, but because of, the quality relationship between researcher and participant (Maxwell, 2007).

Data Analysis

The principal methods of data analysis employed in this study, and outlined below, are consistent with qualitative case study approaches (Creswell, 2013). Raw data was reduced into manageable chunks (Miles & Huberman, 1994) and important information extracted and organized (Flick & Gibbs, 2007) to make sense of what I experienced and observed in the data gathering process (Glesne, 2011). As Creswell (2013) recommended, I analyzed in three phases: (1) data was organized; (2) data was reduced into manageable chunks, categories, and themes through a coding process; and (3), data has been presented through narrative discussion, figures, and tables, below.

In this qualitative research, data collection and data analysis were not exclusive steps, but occurred simultaneously, as suggested by Baxter and Jack (2008). I analyzed data concurrently while collecting it, allowing me to continuously reflect on the data, organize it, look for discoveries, and shape the study as it unfolded (Glesne, 2011; Patton, 2002). Bogden and Biklen (1998) noted that because the researcher in qualitative studies

is a key instrument in the study, responding and adapting to data within the context of the study, he is able to observe nonverbal cues, process data in the moment, summarize, and explore, while conducting the study itself.

Organizing the data. A lap-top computer served as a “fieldwork notebook” (Glesne, 2011, p. 213) for organizing data for analysis and to keep track of interview questions, participant-observation notes, interview notes, and my own personal reflections. Additionally, all documents provided by the school district were entered in Microsoft Excel files and organized initially by school name, administrator, or teacher. Data was analyzed while documents were being reviewed and later as categories were formulated. Likewise, as Morgan (1997) recommended, emerging data was analyzed during and after the interviews, and as field notes, comments, and other observations were added to transcripts (Wolcott, 1994). Demographic information was added to the Excel files to ensure the participants were attached to the correct locations and documents. Interviews were recorded digitally using Garage Band, an application available on most lap-top computers, and transcriptions were made directly into the Excel file, first chronologically to maintain context, and second, by participant.

I read and re-read transcripts and documents, allowing me to continuously sort through it for information important to the study and to thoroughly know the data as a stepping off point for beginning to analyze it (Marshall & Rossman, 2006). Notes were added with each reading, in preparation for initial codes, categories, and themes to be developed (Wolcott, 1994).

Generating categories and themes. Significant data were extracted from

documents and interviews to identify themes within the data. Interview transcripts and notes from pertinent documents were coded and analyzed using a coding schema, as recommended by Khan and VanWynsberghe (2008) and Stake (2005), similar to the schema shown in Table 2. Codes were reduced to categories and themes, and examined for similarities and differences, providing me with themes from which to create meaningful interpretations, draw conclusions, and make inferences and generalizations (Creswell, 2013; Stake, 1995; Yin, 2010).

Beginning with coding of the interview transcript, key words in participants' comments were identified and categorized in terms of what the comments could tell me about participants' perceptions of the evaluation process as a whole. For example, on the left of the coding schema are the questions I asked in the interviews. Participants' comments were added directly from the transcription to the schema into the columns labeled as research question 1, 2, or 3, depending on the question the response was

Table 2

Coding Schema Teacher Focus Group

Interview questions	Research question 1	Research question 2	Research question 3	Other themes
Does <i>Evaluate Davis</i> improve the school?				
Has being observed more been beneficial to teachers and their students?				
Are parent and student surveys valid and helpful forms of feedback for teachers?				
Has increased feedback helped teachers improve instruction?				
Are teachers hired within the last 5 years more likely to improve based on <i>Evaluate Davis</i> feedback than more experienced teachers?				
Is <i>Evaluate Davis</i> a good tool for improving instruction and increasing student learning?				

deemed to fit. The columns expanded as needed, which was helpful in terms of analyzing because it provided a graphic representation of where most of the comments were (Miles & Huberman, 1994). Teachers' names were added so that comments could be easily linked with document data. The schema lent itself to easy development of a story line that naturally weaved in participants' comments and the researcher's insights (Hollway & Jefferson, 2000) to data from current literature and document reviews.

Coding of interview transcripts was followed up with coding of document data by adding quotes from documents to the schema under the headings labeled for the research questions. Codes were compared, linked, and reduced to themes and categories that were connected to themes found in the literature, and finally, I added personal insights and opinions to set the stage for interpretation.

Interpretation. In this phase of the study, data was further reduced, from themes and categories, to the broad meaning of the data, or in other words, what I learned from the data that may be of value to consumers of the study (Wolcott, 1994). Following the recommendation of Gubrium and Holstein (2002), relevant literature linked current theory and practice to the themes and categories identified in the interview transcripts and document reviews.

Interpretation began, as described by Creswell (2013), with coding the data, developing themes from the codes, and organizing the themes to make sense of the data. The categories and themes generated provided the framework for interpretation through thematic analysis and led to naturalistic generalizations of what I learned, as recommended by Lincoln and Guba (1995). Melrose (2009) described naturalistic

generalization as a process through which others develop insight by contemplating the findings presented in case studies, recognizing the similarities that resonate with their own experiences, and reflect how their situations may be similar enough to the researched cases to merit generalizations.

Trustworthiness. As advised by qualitative researchers (Holloway & Jefferson, 2000; Miles and Huberman, 1994) I used multiple sources of data and “multiple perceptions” (Stake, 2008, p. 113) of participants to triangulate evidence and information to improve the trustworthiness of the interpretations generated. Following Glesne’s (2011) recommendations to enhance the trustworthiness of qualitative research, the researcher: (a) examined multiple forms of data (primarily from document reviews and focus group interviews) and engaged more than one type of respondent (teachers, who are evaluated, and administrators, who evaluate); (b) “consciously and continuously searched for negative cases” in the data; (c) was keenly aware of personal biases and “theoretical predispositions” of the participants and the researcher; (d) spent time with the data, at the research site, and developing relationships with participants; and (e) shared interpretations with respondents to verify their accuracy (p. 211).

In addition to focus group interviews and document reviews, my views as the researcher-participant, and the extensive literature review, were used to verify evidence and interpretations, providing strength and support for the findings of the study (Creswell, 2008). As recommended by Wolcott (1994), the themes identified through analysis of documents and focus group interviews were linked with the themes that became apparent in the literature review and the perspectives of the researcher. The final

analysis has created an in-depth, trustworthy interpretation of the common themes found within this case.

The role of the researcher. Glesne (2011) explained the role of the qualitative researcher/author as an interpreter who adds to the data his own “experience, knowledge, [and] theoretical dispositions” (p. 220). Melrose (2009) described the role of the researcher as an interpreter who makes inferences about the relationship between factors in the data that make it applicable to others.

As the researcher in this study, I have an expert perspective as an administrator with 30 years of experience evaluating teachers and evaluating other administrators’ implementation of the evaluation system being studied. Furthermore, I have a unique knowledge of the new evaluation process implemented in the schools being studied due to participation in the formation and implementation phases as a committee member at the district level, with parents, teachers, and administrators. Therefore, I am well qualified to interpret and make inferences about the relationship between factors in the data that make it applicable to others. For this reason, my perspective has been appropriately included, along with the perspectives of other participants, as advised by Patton (2002).

Presenting the data. Data should be organized for presentation in ways that invite the reader to hypothesize and think deeply about meanings and interpretations (Wolcott, 2001). Glesne (2011) recommended several methods of presentation, including a thematic approach in which the writer addresses each theme individually and considers how each can inform further inquiry and practice. Findings, interpretations, and

generalizations of this study are presented through a thematic approach to narrative, supplemented with tables, graphs, and charts (Miles & Huberman, 1994; Wolcott, 1994). The outcome is an organized presentation that provides those charged with using and implementing teacher evaluation systems the strategies and tools that may aid the evaluation process and lead to better teaching and increased learning.

Summary

This chapter described the qualitative research methodology I used for this study, including the case-study design, research questions, data collection process, and analysis. Chapter four explains how and why the evaluation tool, *Evaluate Davis*, used in the Davis School District, in Davis County, Utah, was selected for this case study, and how the subsequent data analysis may be important to implementation of similar evaluation tools. *Evaluate Davis* is described in greater detail along with descriptive data about the participants. The data reduction process is reviewed and the findings are presented as they relate to the research questions. An analysis of the data is also presented that includes four broad themes: (1) Strong agreement is apparent between teachers, administrators, and district level documents that the most important purpose of evaluation is formative: for teacher growth and improvement of teaching; (2) Attitudes, perceptions, and beliefs of all participants in the process are crucial to the formative success of the evaluation instrument; (3) The summative aim of evaluation impacts participants' attitudes regarding the process; and (4) Despite the proclaimed formative aim of evaluation, the attitude that it is a perfunctory ritual or hoop to jump through persists.

CHAPTER IV

RESULTS

I conducted this research with the goal of acquiring greater understanding of teachers' and principals' attitudes, beliefs, and perceptions regarding a teacher evaluation tool, *Evaluate Davis*, and of gaining insight into how their attitudes and beliefs impact the implementation of the evaluation. Insights gained from this study may be used to create an approach for implementation of *Evaluate Davis*, and similar evaluation tools, that may improve teaching, increase student learning, and improve schools. Related questions that guided this study are: What are administrators' and teachers' perceptions of the role *Evaluate Davis* plays in improving student learning gains? What are the perceptions of teachers and administrators respecting how *Evaluate Davis* is used to improve teaching and learning? To what extent do teacher and administrator attitudes regarding the efficacy of *Evaluate Davis* impact its usefulness in improving teaching and learning?

In this qualitative case study, I utilized focus group interviews and document review to collect data regarding principals' and teachers' attitudes, beliefs, and perceptions respecting the implementation and efficacy of the evaluation system, *Evaluate Davis*, used in the Davis School District. Two 45-minute focus group interviews were conducted, one with six teachers and one with six administrators. During the interviews, participants were initially asked questions from the interview protocol included in Appendix B. However, the intent of the questions was to generate controlled discussions through which I sought to gain an understanding of participants' experience with *Evaluate Davis* and to obtain information about their perceptions of the influence of

the evaluation tool on teaching and student learning. Additional questions were asked about specific components of Evaluate Davis: classroom observations, parent and student surveys, student growth, self-assessment and reflection, goal setting, and professional development plans. My role as participant-researcher facilitated a relationship of trust and openness with participants and a deeper, more thorough understanding of their beliefs, attitudes, and perceptions.

I examined two types of district-generated documents: First, documents relating to training teachers and administrators regarding how to implement the evaluation tool with fidelity, and second, teacher observation reports, available to administrators in Davis School District, that provide information regarding the fidelity with which *Evaluate Davis* has been implemented in schools by administrators and teachers. Training documents included district-generated Power Point presentations for new teachers and Power Point presentations prepared for school administrators for mandatory yearly training with faculty and staff. Included with the Power Point presentations were leader notes to be used as a script for administrators as they train their faculty. Additionally, the district provides rubrics for each of the five standards measured in the evaluation: plan, teach, check, environment, and professionalism. Examples of each document are included in the appendices.

The following four observation reports were examined in the document review. (1) *Evaluate Davis Observation Item Report*. This report compares the ratings of the school to the ratings of the district, across the thirty-four indicators in the observation protocol. (2) *Evaluate Davis Observer Item Report*. This report lists all of the school

administrators in the district, by school, and shows the percent of 1s, 2s, 3s, and 4s each administrator gave, the percent total for each school, and how many observations each administrator completed and the total completed for each school. (3) *Evaluate Davis Observation Completion Report*. This report lists each school in the district and shows the percent of teachers with at least one, two, or three observations, and the average observation rating. (4) *Classroom Observation Detail Report*, which lists each teacher in the school and gives the number of times they were observed along with the percent of ones, twos, threes, and fours given.

Also included in this study is a literature review that contains relevant research and information about the characteristics and components of past and current teacher evaluation systems, the iterative improvements to the evaluation and supervision of teachers over the long history of public education in the U.S., and the perceived impact on teacher growth, student learning, and school improvement evaluation has had. The literature review provides the historical context in which the current evaluation was studied, the theoretical framework the study followed, and a detailed description of the characteristics of contemporary evaluation systems—the role of the principal, current issues, including the often-conflicting summative and formative roles of evaluation, the traditionally poor perceptions of teacher evaluation held by many teachers and administrators, and the difficulty in defining quality teaching and connecting it to student learning. Finally, the literature review discusses current expectations for teacher evaluation to impact teacher quality and student learning.

The theoretical framework undergirding this qualitative study includes Moore's

Theory of Transactional Distance, which describes the difference in expectations for the evaluation process between those who use the evaluation tool and the policy makers who require it. A second piece to the theoretical framework is Roger's *Diffusion of Innovations* theory, which clarifies the stages in which innovations, such as a new teacher evaluation tool, are advanced within a social system, such as a school district or the broader educational community. Finally, a third piece is the Marzano Evaluation Model which is built on three theoretical precepts: (a) Teachers who are more skilled and expert produce greater gains in student achievement; (b) The evaluation process has the potential to help teachers to incrementally improve in skill and expertise; and (c) The methods for improving teacher skill and expertise are clear. In this chapter, I present a summary and analysis of the data I gathered through the focus group interviews and the document review.

Summary of Data

In this qualitative case study, I focused on a small sample of teachers and administrators in the Davis School District who have completed at least one evaluation cycle in the evaluation tool, *Evaluate Davis*. Davis School District is a large school district with headquarters about twenty miles north of Salt Lake City. The district covers the whole of Davis County which is a narrow strip of land, about five miles wide, that runs north from Salt Lake County, on the south, approximately 30 miles to the Weber County line on the north. The west border is The Great Salt Lake and the east border is the Wasatch Mountain Range. Eight high schools, sixteen junior high schools, sixty-two

elementary schools, and 2,800 teachers and administrators serve approximately 78,000 students in Davis School District.

Evaluate Davis is typical of the new, more rigorous evaluation systems researchers and policy makers in general are looking to, and that are being implemented in most states in the U.S. today (Donaldson, 2016). The *Evaluate Davis* manual explains that the evaluation process is expected to improve teaching and learning in this way.

Quality licensed educators are essential to effective teaching and learning. In the Evaluate Davis system, educator-created performance standards aligned with the Utah Effective Teaching Standards outline the performance expected of licensed educators. Multiple components in the evaluation system assess the level of performance on each of these standards and indicators and provide information to supervisors as they rate licensed educators on end-of-year summative evaluation reports. The data from the Evaluate Davis components helps teachers reflect, analyze progress, and set goals for the year. (Davis School District, 2014, p. 1)

Evaluate Davis is based on the following principle.

An effective evaluation system recognizes areas of strength and identifies areas for improvement by using multiple measures to assess performance based on standards and performance indicators. The process provides opportunities for continuous professional growth, feedback, and performance assistance, to ensure high quality staff and *learning first* for each student. (Davis School District, 2014, p. 1)

Evaluate Davis requires teachers to engage with six components: self-assessment, observations, a professional growth plan, student growth measures, parent and student surveys, and a summative evaluation. The process begins new each year, but is intended to be ongoing throughout the year, following a cycle that begins with a review of data that includes administrators' previous observations, parent and student surveys, and student achievement data. Reflection, self-assessment, and goal setting follow the data review. Teachers are required to complete a goal setting form that includes one

performance goal and two student achievement goals. Teachers then create a professional development plan aligned with their goals. An administrator reviews the self-assessment, goals, and professional development plan with the teacher early in the school year to assure alignment and provide support and encouragement. Administrators meet mid-year individually with each teacher to check progress toward goals and to review data from new observations and surveys. A summative evaluation is completed mid-year, as well, for provisional teachers. In a final meeting at the conclusion of the school year, data is reviewed again and a summative evaluation completed (Davis School District, 2014).

Teachers are assessed on five domains: Plan, Teach, Check, Environment, and Professionalism. Each domain is comprised of standards, which in turn include specific performance indicators. Observations, surveys, and self-assessments are aligned with the standards and indicators in each domain. Provisional teachers are observed four times each year and career teachers are observed three times. Observations are recorded electronically and include comments and a checklist rating system based on a four-point scale: 1 = not observed, 2 = somewhat evident, 3 = evident, and 4 = extraordinary. Observations are informal, unscheduled, and last at least twenty minutes. Principals are encouraged, but not required, to meet with teachers after each observation (Davis School District, 2014). The online addresses for the Effective Utah Teaching Standards and Evaluate Davis manuals may be found in Appendix F.

In this qualitative case study, I focused on the experiences, beliefs, and attitudes of six teachers and six administrators with the evaluation tool *Evaluate Davis*. The participants' pseudonym, age, the grade levels they teach, their years of experience, and

their primary subject taught are displayed in Tables 3 and 4.

Data Reduction Process

I reduced data into manageable chunks, categories, and themes through a coding process in the following stages.

1. Digital voice recordings and subsequent transcriptions allowed me to secure participants' comments verbatim. Interview transcripts were verified by each participant. Interview transcripts and documents were read and re-read, and relevant information was extracted.

Table 3

Teacher Demographic Information

Pseudonym	Age	Years	Level	Subject
Loraine	60	20	Jr. High	Art/English
Christy	43	13	Elementary	5 th grade
Tom	36	8	High School	Counselor
Karen	50	27	Jr. High	Math
Nadine	54	23	Jr. High	Social Studies
Tara	44	14	Elementary	Resource

Table 4

Administrator Demographic Information

Pseudonym	Age	Years as administrator	Level
Marshall	50	16	Jr. High
Seth	56	16	Elementary
Dan	55	17	Elementary
Brad	46	18	Jr. High
Trent	33	4	High School
Jill	55	20	Jr. High/Elementary

2. Relevant information, common vocabulary, experiences, and ideas shared in the interviews were identified and sorted by research question using a simple coding schema (see Appendix E). Statements and words from documents were isolated by research question, as well, and added to the coding schema.
3. Codes were assigned to every piece of relevant information, sorted into themes, and related to the research questions, by category: documents, teacher interviews, and administrator interviews.

Findings Related to Research Questions

This section summarizes the findings and themes sifted from the data gathered through document reviews and teacher and administrator focus group interviews. The findings are organized and reported by data-gathering method, according to the three research questions. A summary includes a cross analysis of the findings from the two interview groups and the document review and a discussion of the ensuing themes. Interview questions were created to evoke rich discussion regarding the research questions guiding the study. Therefore, many of the participants' comments and many of the statements taken from the documents I reviewed provide insight into more than one research question. Consequently, in the ensuing dialogue, a statement or comment may be repeated several times as it relates to the question being addressed, and as it relates to others' comments, other questions, or district documents.

Document Review Results

The documents I reviewed for this study included training power points, leader notes, leadership training rubrics, and the four district-generated observation reports, as described in Methodology Chapter III. The results for each research question are delineated below according to the research questions.

What are administrators' and teachers' perceptions of the role Evaluate Davis plays in improving student learning gains? In the documents I reviewed there is strong evidence regarding perceptions that specific teacher behaviors impact student learning and that *Evaluate Davis* was designed with the belief that the evaluation process has the potential to improve teaching and increase student learning.

For example, the training materials for Evaluate Davis include Leadership Training Rubrics for each of the five domains: Plan, Teach, Check, Environment, and Professionalism. Teachers are observed and rated on these five domains throughout the school year using these rubrics. Each of the five leadership training rubrics have five columns (see Appendix D). The first column is labeled "Standard and Indicator" and describes a specific teacher behavior such as, "Constructs relevant, meaningful learning experiences which meet individual learning needs." The next column, "Classroom Observation Measure" describes what an observer would see the teacher do in the classroom. For this standard and indicator, the rubric states, "Activities help learners master content or skills." The fourth column, "What Evidence Might I See?" explains specifically what the observer should expect to see students doing. For this standard, for example: Most students complete the assignment in approximately the same amount of time; some students have modified assignments; students share comments on how the concept relates to their interests or occurrences in their lives.

Numerous examples are found within the leadership training rubrics in which the intent to improve teaching with the hoped-for result of greater student learning is apparent. Statements like "Constructs relevant, meaningful learning experiences which

meet individual learning needs” and “Activities help learners master content or skills” are prevalent within the training materials and lead the reader to the conclusion that the evaluation process is planned to improve teaching.

The following excerpts from the Leader Notes in the Power Point for training new educators regarding *Evaluate Davis* show clear evidence regarding the emphasis on student growth as an indicator of teacher effectiveness: “Utah state law requires all certified employees to include evidence of student growth and performance as part of their annual evaluation”; “Each of us, as Davis District employees, impact student learning”; and “An important new requirement (of *Evaluate Davis*) is that (teacher) goals be connected to formal student assessment in order to demonstrate growth or achievement.”

The following slides from the 2016 Evaluate Davis Training Power Point for All Faculty demonstrate a clear expectation that teacher improvement is the anticipated result of implementation of Evaluate Davis.

- *It's All About Growth*
- Slide 3 shows the teacher growth cycle: *Review Data, Self-Assess, Set Goals, Plan Professional Development, Implement Your Plan, and Repeat*
- *Highly effective teachers care about their performance and continually strive to learn, improve, and grow professionally.*

Two important columns in the Leadership Training Rubrics, “Why is it important?” and “Red Flags” point to a belief that what teachers do impacts how much students learn and that there is clearly a connection between certain teacher behaviors and what an evaluator will observe students doing. For example, for the standard/indicator T.1 *Articulates learning goals, content, instructions, and expectations clearly*, the

following statement appears in the “Why Is It Important?” column: *Teachers who articulate what students are supposed to learn and who pre-identify the essential knowledge and skills students will take away from the lesson increase student learning.* In the Red Flags column, references are made to both teacher behaviors and student behaviors. For example:

- When a student is asked what they are currently studying, the answer is something like, “Chapter 6,” or “We have to cut this out and glue it here”
- Posted objective is not written in student friendly language
- Objective is posted on the board but teacher does not refer to it
- Teacher refers to objective only at the beginning of the lesson
- Teacher meanders through lesson and you are unable to tell how content, learning activity, etc., is related
- Student cannot answer the question, “Why are you learning this?” or “Why is this important for you to know?” Some students are not listening to instructions.
- There is confusion among students about page numbers, assignments, missing handouts, etc.
- Students begin work before teacher is finished giving instructions and then ask questions about requirements
- Students have many procedural questions: “I have finished. What should I do now?” “Where do I turn this in?” “When is this due?”
- Students cannot answer the question, “What are you supposed to do?”

Similar examples are found in each rubric.

Examples appear in other training materials as well. For instance, the script included with the Power Point to train new teachers regarding *Evaluate Davis*, The Evaluate Davis for New Educators - Leader Notes, includes a number of references to the

connection between teacher performance and student learning:

- Each of us, as Davis District employees, impact student learning... teach, and provide safe and effective learning environments for students. Our mission of Learning First is achieved through highly effective employees who care about their performance and continually strive to learn, improve, and grow professionally.
- We believe the evaluation system provides employees with the opportunity to engage with administrators and colleagues in meaningful conversations about the practice of teaching and learning.
- You will get the most “bang for your buck” in your goal setting, if you can identify areas for improvement rather than focusing on just one performance indicator.
- Each year you will be setting three goals. Two will be focused on student academic growth/achievement and one will be focused on an Evaluate Davis performance indicator.
- ...we would like at least one Performance Goal to be individual and reflective of what the teacher’s data indicates an area for improvement might be.
- ...the observation doesn’t give us ALL the information about what goes on in your classroom all the time. But, it does give you some good information and, over time, you’ll be able to see patterns in your observation data, providing you with insight regarding your strengths and areas for improvement.
- ...it is important to use technology often and to use it in ways that really impact student learning.
- Please be proactive in self-evaluating and improving your performance. Review your student achievement, observation, and survey data carefully.
- We begin each year by reviewing our data, reflecting on our practice, and setting goals to help us continuously grow and improve. (Davis School District, 2014)

What are the perceptions of teachers and administrators respecting how Evaluate

Davis is used to improve teaching and learning? Each year the district requires school administrators to train their faculty regarding *Evaluate Davis* procedures and policies.

They provide a Power Point presentation and leader notes to assure consistency from

school to school. The most recent training materials I reviewed for this study, *2016 All Faculty Evaluate Davis MicroPD*, describes how administrators should use the evaluation tool to improve teaching. For example, slide 3 describes the yearly teacher growth cycle administrators are expected to guide teachers through: review data, self-assess, set goals, plan professional development, implement the plan, repeat. Slide 5 describes the process: “Teachers and administrators use the Observation Item Report to gather data about needs for personal and faculty professional development.”

Slide 7 advises teachers’ self-assessment to be “Honest, Thoughtful, Reflective.” The two subsequent slides include the following quotes: “Highly effective teachers care about their performance and continually strive to learn, improve, and grow professionally” and “By three methods we may learn wisdom: First, by reflection, which is noblest; second, by imitation, which is easiest; and third by experience, which is the bitterest” (Attributed to Confucius). Administrators are advised to use *Evaluate Davis* to train teachers to be reflective, honest, and thoughtful in their self-assessment.

From the *Leader Notes* for the *2016 All Faculty Evaluate Davis MicroPD*, the following quotes point to the expectation that the evaluation process will motivate teachers to take steps toward improvement that increases student learning.

- In terms of pedagogy, the most important part of the academic goal is the setting of action steps which the teacher will take to achieve the learning targets they’ve set.
- The Action Steps portion of the Academic Goal template is where you will describe the steps you plan to take in order to help your students reach the learning targets you’ve set for them. You can develop this plan alone, but you may want to work together with other teachers, for example on a grade-level team or in a PLC - to map out action steps together.
- An important new requirement is that goals be connected to formal student

assessment in order to demonstrate growth or achievement.

However, the *Leader Notes* for slide 5, which references the goal setting step of the growth cycle, point to a view of the process as less about change and growth and more about jumping through the evaluation hoop: “Most of the goals fit on a single page and describe elements of best practice instruction and assessment you are most likely already doing.” Likewise, the following quotes from *Evaluate Davis Training for New Educators - Leader Notes* seem to show the same contradiction. On the one hand, teachers are told “We begin each year by reviewing our data, reflecting on our practice, and setting goals to help us continuously grow and improve.” Yet administrators are trained to reassure teachers regarding the reliability of the ratings rather than focus on the goal of growth and improvement, as indicated in the same training document.

Discuss the training you’ve participated in with other administrators to increase inter-rater consistency and any inter-rater reliability work you have done with your administrative team at your school. Most teachers don’t know about this and find it both interesting and reassuring to know that administrators are working hard to rate fairly and consistently with other administrators across the district.... Make sure educators understand that the observation is NOT the evaluation.

Such inconsistencies in the focus of training materials may lead both teachers and school administrators to be confused about how the evaluation tool is to be used: for formative purposes of growth or for summative purposes based on ratings and scores. The training materials show a concern with how teachers and administrators perceive the evaluation tool is to be used. For example, the following slide for teachers: “Survey data enables us to see ourselves as others see us. Even though that can seem a little scary, it’s good information.” The *Leader Notes for Administrators* for the same slide say “Teachers may ask about the parent response rate”—a clear concern for the validity of the data

rather than what the teacher can learn from it to improve. Likely, those who developed the evaluation recognized that teachers' concern over the summative nature of the evaluation would focus their attention on the validity and fairness of the scores instead of the formative features.

During the development of Evaluate Davis, teacher focus groups were used to produce rubrics for each element teachers would be evaluated on - Plan, Teach, Check, Environment, and Professionalism. In the *Leadership Training PROFESSIONALISM Rubric*, the importance of teachers' attitudes is noted in the following admonition to evaluators:

In every focus group, teachers indicated they wanted to work with positive people in a positive environment. Negativity not only affects the person who is negative, it also brings down anyone around them.

Teachers who are pro-active, rather than reactionary, are an asset to any school; they help to identify ways to moderate difficulties associated with change and move forward in a positive way. Teachers who grumble and complain create a negative working environment for all. In addition, focusing on the negative decreases the ability of others to move forward efficiently and takes the focus off of student learning. (Davis School District, 2014)

The actual teacher behaviors and attitudes evaluators look for are also delineated in the rubric.

- Embraces, rather than tolerates, initiatives from school and district administration
- Looks for ways to improve and creates effective and productive strategies and action steps for implementing change
- Helps others focus on the positive aspects of the change or challenge facing their grade level, program, or school community
- Uses phrases such as, "I'm wondering if we could...", "What if we tried...", or "Something we might want to think about is..." instead of "That won't work...", "We can't...", "There's no way...", etc.

- Expresses concerns to the appropriate people in the appropriate way with alternative, positive suggestions for improving change process, program implementation, etc.
- Engages in morale building and boosting
- Resists changes in school programs and policies, requirements
- Grumbles about initiatives, publicly or privately
- Talks negatively about the change, the administration, or others involved in the change
- Engages in arguments rather than positive, productive dialogue
- Focuses on negative aspects of program or change rather than imagining the possibilities

The concern with teachers' attitudes in the work place transfers to concerns for how teachers perceive and react to the rating scale, as expressed in the following excerpt from *Training for New Educators - Leader Notes*: "Remember, being *Effective* is great! Some people think of *Effective* as a C (grade), but it is not! *Effective* is an A. A *Highly Effective* rating means you are doing something really incredible - just hitting that indicator right out of the ball park!" In the same document, school leaders are advised, "Teachers may ask about parent response rate" recognizing teachers' concerns over summative ratings using parent and student surveys.

In another training document, *Things to Think About: Validity and Reliability*, the question is posed, "If my supervisor is stricter than others, will my *Evaluate Davis* ratings be artificially lowered?" This question implies a focus on teachers' possible concern for ratings rather than feedback, evidence of concern on the developers' part that the summative role of the evaluation tool may keep teachers from embracing the formative attributes. The same document answers the question thus: "The key is to conduct

sufficient training and analysis to give confidence that concerns with inter-rater reliability will not cause errors in high-stakes employment decisions.” Again, the concern on developers’ part for teachers’ “confidence” in the rating process and how the evaluation will be used as well as fears regarding summative “high-stakes employment decisions.”

The following email a district-level administrator sent, encouraging school administrators to complete observations by an arbitrary date, may reinforce the belief some administrators might harbor that the evaluation process is a hoop to jump through to be in compliance with policy regulations.

Reminder-- In looking over the teacher observations you all have performed; I thank you for staying up with these. A couple of schools need to recommit to get these all completed by the end of 3rd term. Oh! That’s today! Principals, would you go over all teacher observations with your team and review who needs to be observed, knowing that all provisional teachers need 4 minimum observations and then set up a plan to complete these. (Personal communication, March 24, 2017)

That email, coupled with the complete absence of any communication from district level sources regarding the quality of school administrators’ evaluations or how effective the process is in helping teachers improve, supports the perception many hold of the perfunctory nature of the evaluation process.

To what extent do teacher and administrator attitudes regarding the efficacy of Evaluate Davis impact its usefulness in improving teaching and learning? A slide in the *2016 All Faculty Evaluate Davis MicroPD* reminds teachers “Remember, observations are feedback; they are not the evaluation.” This reminder alludes to the designers’ recognition that teachers’ perception of the evaluation’s efficacy is important with respect to the expected impact it may have on teacher improvement. Another slide in the same

Power Point presentation states, “Highly effective employees care about their performance and continually strive to learn, improve, and grow professionally.” This slide suggests the need for teachers to “care about their performance” and implies that without that attitude, they will not “learn, improve, and grow professionally.”

In each of the five *Leadership Training Rubrics* for the elements Plan, Teach, Check, Environment, and Professionalism, there is a column with the heading “Why is it Important?” referring to the standards and indicators. The inference is that the outcomes of the evaluation process are impacted by teachers’ and administrators’ beliefs about the importance of the standards and indicators.

In the *Evaluate Davis Training for New Educators - Leader Notes* there is additional reference to the impact of believing in the system in the following extract: “We believe the evaluation system provides employees with the opportunity to engage with administrators and colleagues in meaningful conversations about the practice of teaching and learning.” Further data from the same document suggests that the designers of the system saw the need to persuade teachers and administrators of the usefulness of the components of *Evaluate Davis*. With respect to Observations:

- Observations allow supervisors to give educators ongoing feedback about their performance throughout the year.
- Observations can give you some good information and, over time, you’ll be able to see patterns in your observation data, providing you with insight regarding your strengths and areas for improvement.
- Make sure educators understand that the observation is NOT the evaluation. The supervisor will use everything they know about the teacher when they complete the end of year summative evaluation. The observations are only a part of that information and provide teachers with feedback on what you actually see or don’t see in that 20 minutes of time. (Davis School District, 2014)

And regarding self-assessment, goal setting, and professional development:

Please be proactive in self-evaluating and improving your performance. Review your student achievement, observation, and survey data carefully. If you are receiving low scores and ratings, work to improve by learning and seeking ideas and input from colleagues or district content supervisors. Take feedback from your supervisor seriously, and work to implement their suggestions for improvement. We want every employee to be successful! (Davis School District, 2014)

Administrator Interview Results

Six school administrators, two elementary and four secondary, participated in an hour-long focus group interview aimed at gathering data regarding the three questions guiding this study: What are administrators' and teachers' perceptions of the role *Evaluate Davis* plays in improving student learning gains? What are the perceptions of teachers and administrators respecting how *Evaluate Davis* is used to improve teaching and learning? To what extent do teacher and administrator attitudes regarding the efficacy of *Evaluate Davis* impact its usefulness in improving teaching and learning? Administrator responses are outlined below as they pertain to each question.

What are administrators' and teachers' perceptions of the role Evaluate Davis plays in improving student learning gains? I asked "Is *Evaluate Davis* helpful in improving the quality of teaching in your school." Jill, a junior high administrator said "I say yes." Seth, another elementary administrator, agreed, with a qualifier: "I think the qualifier is being able to take the time to get back with teachers and to do a good follow up...if you're doing it for the purpose of helping teachers, then it needs to have the time, you need the administrative time to really sit down with teachers and work with them."

Brad, a junior high administrator, added "I think that just the fact there is an

evaluation tool, it helps with accountability. The thirty-four elements to focus on increases specificity of feedback.” A couple of administrators said they believe *Evaluate Davis* is much better (than the previous tool) and a step in the right direction. Trent, a high school assistant principal, said “it is a huge increase in effectiveness” compared to “what we were using...Teachers can look at what is important and find resources to improve. Now we can differentiate.... Much better than what we had in finding what teachers are doing well and where they need help.... I really believe it’s having an impact.” Brad also noted “We come in multiple times each year...at varying times, without warning...and I think that really lends to teachers preparing to meet those thirty-four criteria on a regular basis...I’ve seen that teachers can improve in the areas you address.... Just the process, I think, is helping. When you add to it the opportunity to sit with them and to improve on it, gosh, I think it’s really making a difference.”

Dan, an elementary principal, offered “...I think there is some good that comes from it, but boy, just a matter of getting through it all. To do it exactly as designed, tough, very tough...it’s so time consuming that it’s a difficult thing...I think it’s a pretty good tool.” He also explained that *Evaluate Davis* is “...good to identify teachers who aren’t doing well and aren’t doing what’s best for students and providing them with the support and help they need so that they can grow in the profession.” Trent added of *Evaluate Davis* “...it’s the best I’ve seen. When I talk to people from other districts about how we use it and the data we get from it, everyone is jealous.”

Trent noted that despite some problems with *Evaluate Davis*, “...it’s still effective because you’re having that conversation” with teachers. He added, if “... you continue

to...have productive conversations, then we're going to see an effective evaluation system...and the whole district will benefit." Dan also added "...it's worth the squeeze...it's time well spent...there's enough right about it that I like it and I can see it making a difference...I think we're seeing some results."

What are the perceptions of teachers and administrators respecting how Evaluate Davis is used to improve teaching and learning? (Summative versus Formative) Seth answered a question about the formative aspect of *Evaluate Davis* this way: "If you're doing it (evaluating teachers) for the purpose of helping teachers improve, then you need...time to sit down and work with teachers. If you're doing it just to get the numbers, then no" it isn't helping teachers improve. That statement illustrates what this question is about: "Are evaluators using the summative nature or the formative aspects of *Evaluate Davis* to improve teaching?" Seth talked about the example of an administrator going into a teacher's classroom to observe: "You're there for twenty minutes and the teacher isn't even aware that you're there. You have some humanity, realizing that everyone has a bad day now and again, and you tell the teacher you'll try again on another day." This comment indicates a concern with the ratings, rather than how the teacher can grow from the feedback; a concern with the summative rather than the formative.

Trent expressed the belief that *Evaluate Davis* "...is good..." because it is "...a system of support to help administrators and teachers to become effective..." Dan added that the most important part of the evaluation "...is where that conversation can take place..." where the administrator shares with the teacher areas of strength and weakness he has observed and then develops a plan: "Here are the areas that I've seen or that the

administrative team has seen that you do really well. Here are some areas that you could maybe work on” and so then helping the educator develop that professional plan, “because none of us are perfect.”

Dan shared the following example of how he perceives evaluation is used at his school to help teachers improve: “This particular teacher was doing the right things, saying the right things, but then not waiting for kids’ attention. It was as simple as saying to her ‘Hey, just wait. Wait for ‘em because you’ve got the lingo. It’s okay to pause.’ Once I gave that to her, that was night and day; she wasn’t talking over kids anymore.” He added “I think there’s some good that comes from it, but boy, just a matter of getting through it all. To do it exactly like designed, uh tough, very tough.” Later in the interview he responded “I think it (evaluation) is helpful from a district standpoint instead of just once a year as a hoop to jump through...we can see ongoing throughout the year and occasionally we’ll give feedback.” Later still, he noted “We look at observation data and say ‘Alright, in the planning section, you know, under a certain criteria, we’re missing that, and we’re seeing that often we’re getting a 1 or a 2...this is an area we need to address.’ So, we take those areas, maybe four or five per year, and in faculty meeting we’re trying to address those. And maybe once a month or so, we’ll throw in an evaluator’s tidbit (in an email) ‘hey just a notice, we’ve noticed that we’ve been really great in this area’ or ‘this is an area that we need to focus on and here are some tools for you. Here’s a couple little bits of advice for you.’”

Brad remarked “Surveys are great because most of them are going to fall within the range of the district average. They’re going to be pretty close. I’m looking for

outliers, I'm looking for someone who is well above the district average to commend them, but if they're getting well below the district average in any particular area then you're like, 'hey there's probably a problem here.'" He also noted, with respect to the self-assessment, "What's difficult there is it has a score tied to it. And we don't want to look at the score, or we don't want them to focus on the score because it's formative. So I'm regularly having the conversation with them 'Until we have a summative evaluation none of this counts as a score, don't worry about that. If there is a discrepancy between what I'm giving you in an observation and you're self-assessment, that's not a problem.' But if it's at a summative and there's a discrepancy, then we need to have a conversation. That's probably a critical conversation that we'll have. The self-assessments are almost always much higher than our assessments are. When I look at a self-assessment, I look for kind of holes or weaknesses that the teacher has. So as you watch them throughout the year, if they have complaints against them by parents, you come back to the self-assessment and you see if those match. In other words, the teacher should be saying something like this, 'I'm assessing myself and I'm doing this and this and this, and maybe I'm not realizing I'm saying a few swear words here and there and kids are complaining.'"

Trent answered, "...it's not a hoop because we're identifying teachers who are doing a good job and hopefully we're acknowledging them for that whether it be an email follow up or recognizing them in a class or a faculty meeting or asking them to do a presentation at faculty meeting about certain standard or indicator. But on the opposite end, it's good to identify teachers who aren't doing well and aren't doing what's best for

students and providing them with the support and help they need so that they can grow in the profession. And again I think it's that discussion piece. How we present it to the faculty, how we follow up, and how much time it takes with each of them. I like it because it gives you more personalization with your employees."

Trent further explained his perceptions of how the evaluation instrument is used to improve teaching this way: "I think the summative, and we'll probably talk more about that, has probably given me the opportunity to have more discussions, in-depth discussions, than I ever had (with the previous evaluation system). I think that for us, at least in my experience since that's huge, I love the data. I like how I can get on Encore right now and compare like the inter-rater reliability with not only me and the district, but me and those that I work with at the school. Last year we had an intern at (the junior high) and we identified that he was getting about 33% 1s, so once we figured that out, we sat down as a team and said 'here's our data, what are we going to do about it?' So then we went back and we started doing observations together. Now he had never been a teacher, so we realized he didn't know what to look for in these areas, so it was helpful to use the data and see, um, just make sure there was some inter-rater reliability at the school. Some teachers were getting the raw end of the deal by having a certain person observing them. Um it's also good to uh I think that well we've used the data from *Evaluate Davis* to do at least one school improvement plan goal every year. Um, that's typically part of one of the professional development pieces that we do with the faculty whether that be differentiation or student engagement, whatever it is, we use the indicators from the previous year so that we can say this is what we want to do next year,

and it's a focus for the faculty. And then your last question, do we individualize it? In my experience, I don't know if it's right or wrong, but the only time I've ever really individualized it was when I am working with a remedial teacher and I just need data to back up because they don't understand why we're putting them on this plan, so it's just one more data piece to say 'look, not only are your survey results not that good, but in comparison with the rest of the district, there are on average 5,000 teachers and you're here and that's concerning and we want this plan to help you get back to where you should be.' But that's the only time I've ever used it. I haven't used it for any high teacher to say 'hey 'atta boy, you're doing a great job. Average is here, you're here.' Maybe I should."

Brad noted that the summative evaluation "helps with accountability and improving teaching." He said "What I also love about Evaluate Davis, we've got 34 elements to focus on. We're not just going in and saying, 'Hey, you did well today. You really connect with kids.' I mean, that's part of the equation, but when you can be very, very specific, that is best." He also suggested that the limited time principals have hinders the formative aspects of Evaluate Davis: "Ideally, you meet with the teacher afterwards (after the observation) and you go through that (the 34 indicators) but the reality is, you've got a limited amount of time." Still, he sees that the summative aspect may have a positive impact on teaching. He said "We come in (observe) multiple times each year at varying times, teachers are not warned beforehand; I think that really lends to teachers preparing to meet those 34 criteria on a regular basis. When you identify 34 criteria that are essential to good teaching, and teachers are focused on it, and you're regularly

evaluating it, just that process, I think is helping. When you add to it the opportunity to sit with them and discuss it...gosh, I really think it is making a difference.”

Trent observed “It’s effective in finding what teachers are doing well and where they need improvement...to use the evaluation system and observation indicators to say ‘these are the areas where you’ve been weak. We’re going to observe you once a week and we want to see you meet these criteria...it’s pretty cut and dry; either we’re seeing it or we’re not and we can identify if it’s effective or not.”

To what extent do teacher and administrator attitudes regarding the efficacy of Evaluate Davis impact its usefulness in improving teaching and learning? After being asked if Evaluate Davis was effective in improving teaching and learning or if it is more of a hoop we go through, Seth answered “If you’re doing it for the purpose of helping teachers improve...then yes. If you’re doing it just to get the numbers, then no.” Later in the interview he remarked with respect to the same question “...if the teacher is open-minded and is willing to take constructive criticism, as compared to one who thinks ‘hey, I know it all and there really isn’t room for you to tell me what I’m doing wrong’...I think if their attitude is ‘I’m willing to listen, I’m willing to do that,’ then I don’t see how they could see it as a hoop.” Following up on this comment, Marshall, a junior high school administrator, added “If they just say ‘yeah, yeah, okay’ and nothing changes, then it is a hoop. ‘I have to do it, I have to listen to the boss telling me what’s going on and get it over with and let me just go back to my room.’” Seth added “And I think it also goes along with their perception and the experience that they have. If you have an administrator who is just constantly on somebody about changing and changing and

changing or doing something that the teacher constantly feels attacked, then I think that they're going to see that more as a hoop or as 'let me just do what I do well.' Um at least in their perception. Again, the teachers that are well functioning, that are open-minded, it could be the way that they were approached with it. So I see it as also part of the way the administrator approaches the individual teachers because every building is going to have those that are the pull back or that push back." Another replied "How we present it to the faculty, how we follow up, and how much time we take with each of them...I think a lot of that comes back to the attitude we as administrators take towards it."

Jill offered this example: "I had a really good success story last year because we looked at our data, presented to the faculty, we talked about some really good stuff, did some staff development with it. We went in to do an observation, and mentioned something that I know that he (the teacher) learned or tried and he came with feedback, he's like the veteran teacher forever, and he came back and said, 'you know, I really learned something and I realized I wasn't doing that as much as I should.' I'm thinking, this (the evaluation process) is making sense. You've been to those teachers that have been teaching for some 30 years. But that was a cool thing that he was paying attention and that he was listening and willing to follow through and knew that I was going to follow through. And I think that a big part of this whole *Evaluate Davis* is the accountability."

When asked how effective the self-assessment and goal-setting components of *Evaluate Davis* are, Trent answered "I think it's just like any other goal setting system, if people take it serious, then yes they are very effective. I think that whether or not they

think it's a hoop is based on our response to it. If we're not giving them any feedback after the observation, then whether good or bad, like if we just go in and do the observation, and there's absolutely no contact after that, they're seeing it as a hoop not only for them, but for us too, and they're just thinking 'oh yea, that's my second one, I have one more, I have two more.' And if we can provide some sort of follow up, it's just like a student with a teacher. If a student does a worksheet and there's absolutely no feedback, the student's going to perceive it as a hoop and the teacher is too. Now if there's feedback, like 'I really like how you did this part of the assignment,' now it's no longer a hoop and there's some validation with it. So, whether or not it's a hoop is reliant upon us."

With respect to the self-assessment component, Dan said "I think the self-assessment is more a measurement of their self-esteem." Marshall added, "The self-assessment is okay, but it doesn't hold very much weight" unless the teacher is willing to acknowledge "I'm bad in this area. I need help." He said, "...the goals are good. The self-assessment? Meh." Seth said of goal setting "I think it also goes along with their (teachers) perception and the experience they have." He shared an example in which a teacher may feel attacked because a principal is "constantly on somebody about changing" and the teacher may see the process as a hoop or dig in their heels and respond with "let me just do what I do well." He said, "It could be the way that they (teachers) are approached with it (goal setting)...the way the individual administrator approaches the individual teacher..."

I asked if there were things about Evaluate Davis that might create the perception

that it is a hoop to jump through. Marshall said, “I think it’s a small hoop to jump through. I think it does help when you go through it and really think about it as you do that. I think it can be a hoop, but it kind of relies on us to make it so it’s not.” Trent added, “Sometimes it can feel like that, (a hoop) but it’s, I think, a mentality shift that we (administrators) have to have to make sure that when we go do these observations and when we do these mid-year evaluations and summative evaluations, that we recognize that this is an opportunity to celebrate success and build them for the future.” He concluded “Yeah, um, just in the remarks I’ve heard and thinking back on a lot of research that I read, most principals and teachers in using the previous evaluation...you would see that they kind of just thought of that teacher evaluation as just sort of this thing you do once a year, get it over with, check the box and it’s done.”

Seth added, “You think of the question ‘Is this just a hoop to go through?’ I think a lot of that comes back to the attitude we as administrators take towards it. Because if you think of the last evaluation system...there were still goals with that and teachers would write goals you know, but how often did the teachers come back and say ‘Um I’ve got my evaluation, and do you have a copy of my goals?’ At the end of the year, they’re sitting in a file somewhere and...they never looked at them. Well, that still happens now. ‘How do I get back into Evaluate Davis to look at these (goals)?’ Okay, that kind of tells you that the teachers are seeing that as kind of a hoop. Or that tells you which teachers are seeing it as a hoop rather than using it as a real tool. So to me that’s an indicator as to who’s really engaged in it and who’s engaged in making their profession better and increasing their ability in their craft.”

Principals also noted that in some cases, teachers' perception that the evaluation is summative, that it is going to impact their status in the profession, influences their self-assessment. For example, Trent commented, "...you have some teachers who are so oblivious to reality that they're giving themselves straight fours (the highest possible scores)."

Even students' and parents' attitudes about the evaluation process have an impact. *Evaluate Davis* requires every teacher to have their students and their parents complete surveys every year. On a secondary level, a student will have either eight or ten teachers, depending on which kind of block schedule their school uses. If a family has just one student in the secondary system, a parent could be asked to complete a survey for up to ten different teachers. Considering that surveys take up to ten minutes to complete; that adds up to well over an hour, which is time parents aren't willing to give up unless they have unusually strong feelings one way or the other about a teacher. The time problem is compounded if a family has more than one student in a secondary school. The same issue comes into play for students who are often shuffled into a computer lab and told to log in and complete surveys for each of their teachers. Trent noted "I think the surveys are incredibly weak at the secondary level...we bring them (students) all together...doing their surveys all at once and you just watch kids click, click, click. Whether it's good or bad, it's not valid."

Teachers' attitudes come into play as they write goals each year to meet the requirements of *Evaluate Davis*. As described above, teachers are required to complete a goal setting form that includes one performance goal and two student achievement goals

for each course they teach. This is very cumbersome for teachers who teach multiple courses. To ease their burden, curriculum supervisors at the district level have developed generic goals for almost every course that have become part of the electronic form teachers complete at the beginning of each school year. Teachers simply fill in the blanks with a number or percentage. Several of the principals noted this in their comments about the goal setting component of *Evaluate Davis*. Trent commented “I think it’s just like any other goal setting system, if people take it serious, then yes, it’s very effective.” He went on to say “I think too often people don’t take it (goal setting) seriously because they don’t believe in what a goal means or their curriculum supervisor has set the goals for them and it hasn’t changed since last year...and there’s not a lot of meat in those goals.”

Trent shared “Even as an administrator, when I’m looking to see the science goal and they’re teaching 7th grade, 7th grade honors, 8th grade, 8th grade honors, whatever the situation is, that’s four sets of goals for that teacher and they’re all the same...The teachers aren’t bought in and quite honestly, I’m not either. Now the goals they’re writing themselves are the ones I really care about.” Marshall admitted, “I don’t do much with goals because like everybody said, goals are just kind of set by the school district and that’s what they put down.” Dan made note of the transition, in just four years, from teachers writing their own goals to district curriculum supervisors writing them and how that change has affected his attitude: “The goals, first year, I read every single one. And if it didn’t make sense, I hit the box to reply and give them some feedback, ‘make this make sense.’ When they were writing them themselves, it was amazing that first year how many goals I just needed them to clarify or rewrite so that it made sense. Then it seems

like year two I was getting these well-written goals that were just cut and pasted, then I just hoped they fit the subject area, you know? So, they would cut and paste a little bit too much. I think my second year reading goals they...included too much and it was too big of a cut and paste. But they were better written. Whomever is writing them at the district is doing better than my teachers.”

Brad shared how he tries to help his teachers have the perception that the evaluation is formative: “What’s difficult there is it has a score tied to it. And we don’t want to look at the score, or we don’t want them to focus on the score because it’s (the evaluation) formative. So I’m regularly having the conversation with them ‘Until we have a summative evaluation none of this counts as a score, don’t worry about that. If there is a discrepancy between what I’m giving you in an observation and you’re self-assessment, that’s not a problem.’”

Teacher Interview Results

What are administrators’ and teachers’ perceptions of the role Evaluate Davis plays in improving student learning gains? I asked the teacher focus group if the evaluation process helped them be better teachers or if students were learning more because teachers are evaluated. Teachers’ responses ranged from “Evaluation doesn’t help me be a better teacher” to “It could. There is a possibility.” Karen, a junior high math teacher for twenty-seven years, said “Not so much because I set easy goals. I do because I think it’s a hoop. I do... I set easy goals because I’ve got a lot to do this year. I had curriculum to write, I had 3 different preps and the last thing I wanted to do is spend time on a goal that I think isn’t important. I do what I do well. I do.” Nadine, a twenty-

three year junior high social studies teacher, added that *Evaluate Davis* could "...help some of the lower teachers get better, or at least get identified...but it mostly bugs seasoned, good teachers..." like her. Christy, a fifth grade teacher for the last 13 years, agreed, "...for me, because I'm so driven, it doesn't really affect me a whole lot. I'm very driven and it's just a hoop, but for the teachers that are not driven...that need help, I think on the whole I would say yeah, I think it's helpful, but for me personally, it doesn't make a lot of difference."

Christy said, "I don't think it (*Evaluate Davis*) helps them (teachers) improve, unless they have training." Tom, who has been a high school counselor for eight years, explained "...because we look at it once a year at the beginning. We all pile into a computer lab, they tell us what to put in and then you don't hear about it again for 365 (he meant 180) days and then all of a sudden everyone's like we've got to get in and do it all again at the busiest times of the year (the end), so it becomes just a complete hurdle that we have to jump over as opposed to a tool to help us get better."

What are the perceptions of teachers and administrators respecting how Evaluate Davis is used to improve teaching and learning? (Summative versus Formative) Teachers agreed on a number of ways the evaluation could be formative and improve teaching. They like having administrators come into their classes to "...see what I do...come spend some time in my classroom." Karen responded "I appreciate the evaluations (observations), yeah like when my administrator comes in, I think those are beneficial, to a point...the most valuable thing I think are the observations, when they actually come in your class and see what you're doing." Tara, a fourteen-year special education teacher

noted, "I got zero observations this year. So how is that supposed to help me improve? I had to beg them to come in and observe. 'Oh, we know you don't need it.' Really?"

All of the teachers agreed, "face-to-face feedback" from evaluators, which includes "suggestions for improvement and solutions to problems" is the most favorable circumstance for teacher evaluation to improve teaching. Specific feedback, especially feedback about what teachers are doing well. As one teacher put it "I like the comments, comments about something you did, the good." When something could be improved, they want to know what to do to be better, not "just tell me what's wrong." Loraine, a twenty-year veteran art teacher responded "I actually would like them to come in and give me some feedback...that's my favorite part of the whole thing because I feel like I do really well and I love it when the principal comes in and says 'You're doing so good!' It's like yay! So that's my favorite part of the whole thing."

Teachers' positive feelings about the feedback they get from evaluators wasn't all about being validated or told they're "doing a good job." As administrators use the electronic rating form, they can leave comments by each indicator if they choose to. Nadine emphatically replied, "I love the comments!" Christy replied, "I do too. I think it adds validity to them and lets you know that that person was really watching and not just filling in holes." Loraine added "And that's the one thing I respond to. Like when I got some comments and went back to one of my advisors and said 'well what did you mean here, here, and here?' Otherwise I don't respond, I don't really care (about the ratings), so it's the one thing that I like. I really enjoy them." Nadine concurred, "They're the most valuable of it. Even just a small comment."

Nadine summarized her opinion that the evaluation process is more summative than formative: “I think if its intended goal is to identify weaker teachers, either to help them or get them out, I think it helps identify them. I don’t think it helps improve them. ...I think it provides documentation to get them out. But if the purpose is really to help teachers, it’s a bigger pain in the butt than I think we get out of it. It is a much bigger hoop to jump through and pain because we have to provide a lot and there’s so many levels. But the observations themselves, I would much rather my principal just come in and out of my class often for a few minutes and observing and knowing what’s going on than coming in officially sitting for 20 minutes and filling out a bunch of dots and then leaving. I...get more out of when they just pop in and out and stay for a few minutes, but I like them in my room.”

Teachers viewed the observation rating scale as mostly summative and disliked the summative evaluation ratings as well. They reported having a “hard time” with the rating scale, often confusing the observation rating scale with the summative evaluation ratings. They have become accustomed to receiving the highest scores possible most of the time, noting that the scores on the observation checklist don’t help because “ones and twos just make them mad and no one gets fours; no one knows the standard to get a four.” Nadine said “...I don’t like the 1, 2, 3, 4 scale...the 4 is supposed to be if you’re really good. We’re not allowed to have those so it’s ridiculous. The district doesn’t want evaluators to give them away.” She added, comparing her evaluations with the previous system to those she receives now using *Evaluate Davis*, “... I went from getting like straight ‘you’re amazingly awesome’ to ‘I can only give you a max of five of those

(fours).’ Can you imagine us teaching our class and saying ‘I’m sorry, I can only give five As because that’s all were allowed to do?’ So it’s artificial and it negates the validity of any of it... I just think that there’s an artificial...hidden agenda in there. So to me I don’t care if you give me twos because there’s no validity in any of it.”

The scores, teachers said, feel like a “reward or punishment... not seeing high scores is disheartening.” I probed a little deeper by asking about administrators’ efforts to help teachers see the formative nature of the evaluation, that the scores just indicate a need to look at that item and consider how to do that better. Even though administrators assured them that the rating scale was for formative purposes, to identify “points of discussion”, teachers were still “worried about the record” and what a 2 means. Teachers reported feeling they had been “dinged” if they got a score less than a 3 on any standard and still believed “more 4s should be given out.” Teachers complained that the ratings from individual observations were the “basis for the summative evaluation” and therefore “threatening.”

While most of the teachers agreed that teacher evaluation “could improve teaching” they were less disposed to believe *Evaluate Davis* was actually being used for that purpose. Furthermore, they were not real sure what the rating scale is or means, even though they have used the same system for four years and their school administrators have supposedly trained them and shown them where to find that information. Nadine asked, “Let’s see...it goes ‘not evident’, umm, is it ‘met standard’, then ‘effective’, and ‘highly effective’?” Christy questioned, “So if you’re not a 3 or 4, you’re a bottom 2? Does your principal sit down with you and make goals for you for the next year? Or is

there some kind of re-teaching that happens for those teachers? Because if there's not, what's the purpose? There's no purpose to that other than to say you're bad." In other words, it is only summative.

For example, Loraine talked about the reluctance of administrators to give fours, the highest possible rating: "So what I did to combat it...is when I do my...self-evaluation, I do all fours, like fine, it's all fours. You're not going to give me all fours? I'll give myself all fours. That's how I think the fours are ridiculous." Nadine added, "Giving me a three or four isn't going to change what I do. No, I'm not going to change my approach for a better score. It's if my kids didn't get it, that's when I change it, not when it's on my paper from someone who is in my room for five minutes." Christy countered, "I get a good number of fours. I also get ones if they're not applicable or I've gotten a few twos over the years. But I'm always a highly effective or effective. So it's always good information..."

Teachers reiterated a key point: they see the summative role of *Evaluate Davis* as "hostile" and said they "need help from people who aren't threatening," referring to administrators. Nadine summarized "I wish there was a way that the evaluation could meet better the two sides (summative and formative) because I really think as teachers, we need feedback...but I don't like the scaling (ratings)...That if it (the feedback) says I'm not meeting the standard, I have an opportunity to fix it; my job's not on the line. People feel threatened that their job is on the line, and in some cases it needs to be, but there needs to be a way to identify my weaknesses and have an opportunity to develop a plan to improve."

Not only did teachers not trust the observation rating scale, they didn't feel the surveys were used formatively either. I asked teachers "How valid do you believe parent and student surveys are in terms of the feedback they provide teachers about their teaching?" Their response was unanimous "Not at all!" Their follow-up statements further indicate a lack of trust in surveys as valid, formative data. Loraine identified her concerns: "It's not fair for them to be anonymous; they need to identify themselves." She said she wants to know "...what kind of students hate me." Otherwise, she said, the survey "...doesn't help me beans." A student who gives poor ratings, she observed, could be "...mad at me that day." She said she would trust the student surveys more if she knew "whether the student is a good student or a poor student." She also worried that the student survey could be used "as a way to get back at teachers." Similarly, Christy added "I think that you...get two different kinds of parents that respond, the ones that love you or the ones that hate you, so it's really just an average."

"Not enough parent surveys are turned in" is a concern Tom expressed for the group. He said, "We had eleven percent of our parents respond last year. That's not very good. We looked it up and only one-tenth of our parents took the time." He added "...if parents are frustrated with something, they're going to just hammer you." Karen, who has two children of her own in high school, explained that she didn't participate in completing surveys for her children's teachers because between the two, they have eighteen teachers.

Surveys have numerous other problems teachers noted. Student surveys have a three-point scale, "always", "sometimes", and "never". Teachers criticized the scale,

pointing to the big gap between “always” and “sometimes.” “We need another option, like ‘frequently,’” they agreed. The timing of surveys isn’t good, as Loraine pointed out: “Students are surveyed too early in the year; they don’t really know anything yet.” She added “students don’t own the survey; there is no accountability for anything they say.” Students don’t take the surveys seriously, as Nadine noted: “And the student ones, at least in our school, they just take a day and...sit them down in one of...the computer labs and they do...” surveys “... for all seven (of their teachers) at a time, and they just...tell us ‘I liked you; I gave you straight yeses’ or...bragged that they hated you, they gave you all nos. It has no meaning.”

I asked teachers if the goal setting process felt like a hoop. They answered emphatically “Yes!” Tara added “The goals for sure. I honestly have no idea if I’ve even looked at those since the beginning of the year.” Karen told of how her school has teachers gather in a computer lab on a Friday early-out day and input all their goals at the same time: “...we do ours on Friday afternoon at 3 o’clock in the lab together. I’m writing stuff down just so I can leave.”

They complained that the goal setting component of *Evaluate Davis* is a “hoop to jump through” because “...we get goals given to us.” *Evaluate Davis* requires teachers to set two student achievement goals and one performance goal. Teachers should look at parent and student surveys, their own self-assessment, and the data from their administrators’ observations to create a performance goal. However, both teachers and administrators said they don’t have time to do it that way, and performance goals are rarely based on data. District curriculum supervisors create the student achievement goals

and teachers set their own targets, action steps, and dates for completion. Teachers questioned the practice of district curriculum coordinators writing goals for teachers: “I don’t know how the district can give us goals. That makes no sense to me. I don’t understand that concept at all” Christy said.

Teachers did not seem to understand the difference between the student achievement goals and the performance goal. Teachers are required to have two student achievement goals for every course they teach, so Karen explained that for every goal the district comes up with for her, “...it was one less I had to come up with...” In fact, Nadine gave the perfect example of how the performance goal should work while complaining about curriculum supervisors creating their student academic goals: “I could still, every year, say ‘this is an area I need to work on or I want to work on.’ I wish that was part of my process instead of making up hoops to jump through. This would be valid for me to say ‘I’ve identified this is a weakness or my principal has identified this is an area I need to work on. I’ve set goals of how to improve in my teaching of that, and I’m going to show growth.’” The fact that teachers are unclear on the difference between the types of goals they are asked to set, and they feel like the process for creating performance goals is new to them or something they are discovering, speaks to the reality that the whole goal setting component has not been taken seriously.

Overall, teachers do not like or trust the summative aspect of *Evaluate Davis*. They indicated in their comments that the ratings are superficial and create animosity toward the evaluation process. They do appreciate face-to-face feedback and comments that lead to improvement.

To what extent do teacher and administrator attitudes regarding the efficacy of Evaluate Davis impact its usefulness in improving teaching and learning? Karen, indicated she sets “easy goals...because I think it’s (evaluation) a hoop.” Clearly, her attitude that the evaluation process is “a hoop” effects how much effort she is willing to invest in the formative goal-setting component. Nadine’s belief that “Evaluate Davis has way too many steps” has narrowed her vision of the usefulness of the tool to the “...most valuable thing I think are the observations, when they actually come in your class and see what you’re doing.” She seems disposed to discount all of the other formative components like self-reflection and assessment, goal setting, and professional development. Nadine noted of the formative aspects “I don’t think it helps (teachers) improve; I think it provides documentation to get them out.” Speaking of the formative components, she said, “...it’s a bigger pain in the butt than we get out of it.” Loraine expressed a similar point of view, saying that the evaluator’s comments are “the one thing that I respond to...Otherwise, I don’t respond (to the ratings). I don’t really care.”

Christy said she believes the principal’s attitude “...can really make a big difference.” She explained “I’ve had a really horrible principal” who she observed was most intent on trying “to find something wrong,” and most recently an “amazing” principal whose approach “totally changed the feeling of the whole school.” Tom had a similar experience that has affected how he approaches the evaluation process. Of his current school administration, he noted “...we don’t get comments and we don’t get observed much and so, at the end of the year, they’ll do our little review and say ‘yeah, it sounds like everything is going great...’ and so it feels a little more detached in the sense

that we don't get a lot of feedback." He said he received a lot more from previous administrators at another school, but now he gets "...a rating and a 'good job.'" Karen also agreed saying that her new principal "actually went through everything...my feedback was quite valuable this time, but in other years it has not been."

Nadine reported that she complained to her principal regarding the rating scale and the reluctance of evaluators to give fours, the highest possible rating. Her principal, she says, told her that he couldn't give more than five fours to a teacher. "It's artificial and negates the validity of any of it" she reported, "So to me, I don't care if you give me twos because there's not validity in any of it." Teachers' frustration with the rating scale undermined how they complete the formative self-reflection and assessment piece, as well. Nadine reported "So what I did to combat it, when I do my own my self-evaluation, I do all fours. Like fine, it's all fours. You're not going to give me all fours? I'll give myself all fours. That's how I think. The fours are ridiculous." Tara added "...they don't want me to put a 4 so I have to put something lower than that because I'm not allowed to put a 4. I'm not allowed to be good at that even though I think I'm good at this. It's interesting to try and manipulate really the system."

Tom said of the lack of fours his administration would give out, "I had an assistant principal explain it to me that you always need to be looking to improve. And so that's your motivation. And I argued that if I'm already good, what about maintaining, and obviously we can always change, but if you're good, you're good and if you're not, you're not. So we had a little disagreement about what it meant, so I gave myself 4s just to spite him."

Nadine concluded, “Giving me a 3 or 4 isn’t going to change what I do. I’m not going to change my approach. It’s if my kids didn’t get it, that’s when I change it; not when it’s on my paper from someone who is in my room for 5 minutes. The feedback that you get from your kids, if they learned the concept or didn’t learn.” She explained her point of view that “...good teachers” have learning “objectives tied to student learning” and “have thought through what they want kids to do and have a purpose and direction for their lesson. They’re monitoring themselves already. I would bet all of us do that; we already adjust.” Clearly, teachers’ negative view of the rating scale impacts their willingness to use other formative components effectively.

Teachers reported they view the goal-setting component of *Evaluate Davis* as “a hoop.” Frustration over having “goals given to us” advanced that perception which was also said by teachers to influence their desire to make changes or achieve their goals. Karen said “...the department gave us goals and said ‘This is how you’re going to measure,’ and then there were so many other things they asked us to do that it was hard to do what they wanted.” She added “I honestly have no idea if I’ve even looked at those (her goals) since the beginning of the year...And, we do ours on Friday afternoon at three o’clock in the lab together. I’m writing stuff down just so I can leave.”

Analysis

Organizing data, especially opinions, perceptions, attitudes, and beliefs, into neat compartments or themes is difficult at best. Educators’ attitudes can be all over the place and sometimes contradictory. Analysis of the data I gathered in the focus group

interviews and document reviews, triangulated with my own experience and the findings from the literature, illuminated four main themes in the attitudes of educators that may influence the effectiveness of teacher evaluations in general and most specifically Evaluate Davis.

Themes

Four broad themes emerged from the literature review and data gathered through the review of documents and the focus group interviews: (1) Strong agreement is apparent between teachers, administrators, and district level documents that the most important purpose of evaluation is formative: for teacher growth and improvement of teaching; (2) Attitudes, perceptions, and beliefs of all participants in the process are crucial to the formative success of the evaluation instrument; (3) The summative aim of evaluation impacts participants' attitudes regarding the process; and (4) Despite the proclaimed formative aim of evaluation, the attitude that it is a perfunctory ritual or hoop to jump through persists.

Included within each of the four broad themes are sub-themes. In this section, I examine the relationships between themes as they pertain to the perceptions, beliefs, and attitudes expressed in the documents and teacher and administrator interviews.

Formative. The documents I reviewed, and the data I gathered from interviews, provide strong evidence that *Evaluate Davis* was designed with the belief that the most important aim of the evaluation process is formative; that it has the potential to improve teaching and subsequently increase student learning. Subthemes include: (1) the construct that specific teacher behaviors can be tied to specific student outcomes; (2) data gathered

through classroom observations, self-reflection, and parent and student surveys may be used to identify strengths and weaknesses in teacher behavior; and (3) the most teacher-valued formative tool in the evaluation process is face-to-face communication between administrators and teachers in which specific feedback and comments are directed at improvement.

Teacher behaviors. Evidence is plentiful in the data that I gathered that those who designed Evaluate Davis, as well as teachers and administrators, believe strongly that certain teacher behaviors are essential to the learning process and that when those teacher behaviors are observed, specific student learning and behavior outcomes will be observed, as well. Numerous examples are found within the Leadership Training Rubrics in which the intent is clearly to improve teaching with the hoped-for result of greater student learning. For example, one column in the rubric, “Classroom Observation Measure,” describes specific teacher behaviors an observer should expect see during a classroom observation, and the next column, “What Evidence Might I See?” describes the consequent student behavior outcomes.

Statements like “Teachers who articulate what students are supposed to learn and who pre-identify the essential knowledge and skills students will take away from the lesson increase student learning,” “Constructs relevant, meaningful learning experiences which meet individual learning needs,” and “Activities help learners master content or skills” are prevalent within the training materials and lead the reader to the conclusion that the evaluation process is planned to improve teaching with the belief that improved teaching impacts student learning. The belief in that construct is also apparent in training

slides that state “It’s All About Growth!” and “Highly effective teachers...strive to learn, improve, and grow professionally.”

The training rubrics include a column, “Red Flags,” which also points to a belief that what teachers do impacts how much students learn and that there is clearly a connection between certain teacher behaviors and what an evaluator will observe students doing. In the “Red Flags” column, certain specific student and teacher behaviors are described which, according the training materials, should indicate to the observer that teaching is sub-standard. Examples include statements like “Teacher refers to objective only at the beginning of the lesson” or “Student cannot answer the question, ‘Why are you learning this?’ or ‘Why is this important for you to know?’”

Administrators expressed their view, also, that the intent of *Evaluate Davis* is formative. Trent, a high school administrator who participated in the focus group interview, expressed the belief that *Evaluate Davis* “...is good...” because it is “...a system of support to help administrators and teachers to become effective...” Seth, an elementary administrator, responded to question about the effectiveness of *Evaluate Davis* with “If you’re doing it for the purpose of helping teachers improve...then yes” it is effective.

Data. Data gathered through the evaluation process, from observations, parent and student surveys, measures of student growth, and self-assessment, may be used for formative or summative purposes. Many of the training documents I reviewed suggest that data is meant to be used formatively. For example, from the Leader Notes in the Power Point for training new educators regarding *Evaluate Davis*:

- ...we would like at least one Performance Goal to be individual and reflective of what the teacher's data indicates an area for improvement might be.
- ...the observation doesn't give us ALL the information about what goes on in your classroom all the time. But, it does give you some good information and, over time, you'll be able to see patterns in your observation data, providing you with insight regarding your strengths and areas for improvement.
- Please be proactive in self-evaluating and improving your performance. Review your student achievement, observation, and survey data carefully.
- We begin each year by reviewing our data, reflecting on our practice, and setting goals to help us continuously grow and improve.

A slide from the *2016 All Faculty Evaluate Davis MicroPD* encourages,

“Teachers and administrators use the Observation Item Report to gather data about needs for personal and faculty professional development.” The administrative interviews also produced a number of comments regarding how data can be used formatively.

Administrators suggested that the principal may use data from the evaluation process and sit down with a teacher and say “Here are the areas that I've seen or that the administrative team has seen that you do really well. Here are some areas that you could maybe work on.” Dan noted that formative process works both for individual teachers and for the whole faculty: “We look at observation data and...” identify areas “...we need to address. We take those areas, maybe four or five per year, and in faculty meeting...” we try “...to address those. And maybe once a month or so, we'll throw in an evaluator's tidbit (in an email) ‘hey just a notice, we've noticed that we've been really great in this area’ or ‘this is an area that we need to focus on and here are some tools for you. Here's a couple little bits of advice for you.’”

Trent also explained how his administrative team uses data formatively: “...it's good to identify teachers who aren't doing well and aren't doing what's best for students

and providing them with the support and help they need so that they can grow in the profession...it's also good to use the data from *Evaluate Davis* to do at least one school improvement plan goal every year. That's typically part of one of the professional development pieces that we do with the faculty." Likewise, he noted "It's effective in finding what teachers are doing well and where they need improvement..." Jill had a similar experience she shared: "I had a really good success story last year because we looked at our data, presented to the faculty, we talked about some really good stuff, did some staff development with it."

Even though administrators seemed positive that data from the evaluation is being used formatively, teachers seemed less enthusiastic. Nadine suggested that *Evaluate Davis* could "...help some of the lower teachers get better, or at least get identified...but it mostly bugs seasoned, good teachers..." She added "...there needs to be a way to identify my weaknesses and have an opportunity to develop a plan to improve." Christy said, "I don't think it (*Evaluate Davis*) helps them (teachers) improve, unless they have (receive) training." Roadblocks for them seemed to be their apparent distrust for the summative uses of the same data, to be discussed in a later section.

Face-to-face feedback. Training materials state "We believe the evaluation system provides employees with the opportunity to engage with administrators and colleagues in meaningful conversations about the practice of teaching and learning." Teachers and administrators are cautioned: "Remember, observations are feedback; they are not the evaluation." The following comments from the *Evaluate Davis Training for New Educators - Leader Notes* allude to the belief that feedback is point of the

observation component of *Evaluate Davis*:

- Observations allow supervisors to give educators ongoing feedback about their performance throughout the year.
- Observations can give you some good information and, over time, you'll be able to see patterns in your observation data, providing you with insight regarding your strengths and areas for improvement.

To the administrators who are training their faculty, the *Leader Notes* admonish “Make sure educators understand that the observation is NOT the evaluation. The supervisor will use everything they know about the teacher when they complete the end of year summative evaluation. The observations are only a part of that information and provide teachers with feedback on what you actually see or don't see in that 20 minutes of time.”

The training documents suggest a clear message: the evaluation process is intended to be formative, to provide feedback to teachers and “the opportunity to engage with administrators and colleagues in meaningful conversations about the practice of teaching and learning.” However, a huge disconnect exists between the training materials and what teachers indicate is the most valuable component of the evaluation process - face-to-face feedback and comments. Nowhere in the training documents is there anything dealing with training administrators in how to give quality feedback or how to find the time, as is noted in administrators' comments, although they do acknowledge the desire to spend more time one-on-one with teachers.

Seth, for example, offered, “...you need the administrative time to really sit down with teachers and work with them.” Dan concurred, “I think there is some good that comes from it (evaluation), but...it's so time consuming that it's a difficult thing...” He

also explained that the most important part of the evaluation "...is where that conversation can take place..." where the administrator shares with the teacher areas of strength and weakness he has observed and then develops a plan for improvement.

Trent also concluded that if "... you continue to...have productive conversations, then we're going to see an effective evaluation system... If we're not giving them any feedback after the observation, then whether good or bad, like if we just go in and do the observation, and there's absolutely no contact after that..." it's not effective. "And if we can provide some sort of follow up," he noted, "it's just like a student with a teacher. If a student does a worksheet and there's absolutely no feedback, the student's going to perceive it as a hoop and the teacher is too. Now if there's feedback, like 'I really like how you did this part of the assignment,' now it's no longer a hoop and there's some validation with it." He added that *Evaluate Davis* "...has probably given me the opportunity to have more discussions, in-depth discussions, than I ever had..." Brad also indicated that "Ideally, you meet with the teacher afterwards (after the observation) and you go through that, but the reality is, you've got a limited amount of time."

The teachers I interviewed saw face-to-face feedback and administrators' comments as the most valuable formative tool in the evaluation process. They said they like having administrators come into their classes to "...see what I do...Come spend some time in my classroom." "I appreciate the (observations), like when my administrator comes in; I think those are beneficial...The most valuable thing I think are the observations, when they actually come in your class and see what you're doing" Nadine explained. They indicated they like specific feedback about what they are doing

well, "...comments about something you did, the good" that includes "suggestions for improvement and solutions to problems." They want to know what to do to be better, not "just tell me what's wrong." Loraine said, "I actually would like them to come in and give me some feedback...that's my favorite part of the whole thing." She acknowledged, "...that's the one thing I respond to...when I got some comments." Nadine said "I love the comments" and the others agreed. "They're the most valuable of it. Even just a small comment" she explained, and added, "I really think as teachers, we need feedback."

Attitudes, perceptions, and beliefs. In the documents I reviewed there are a number of statements that imply an awareness of the importance of teachers' and administrators' attitudes toward certain components of the evaluation process. For example, "Our mission...is achieved through highly effective employees who *care* about their performance and continually strive to learn, improve, and grow professionally;" "Survey data enables us to see ourselves as others see us. Even though that can seem a little *scary*, it's good information;" or "Teachers may ask about the parent response rate" which acknowledges that teachers may have negative perceptions regarding parent surveys.

Many of the comments are aimed at teachers' perceptions regarding summative characteristics of *Evaluate Davis*. In the training document *Things to Think About: Validity and Reliability*, the question is posed, "If my supervisor is stricter than others, will my *Evaluate Davis* ratings be artificially lowered?" This question implies a focus on teachers' possible concern for ratings rather than feedback, evidence of concern on the developers' part that the summative role of the evaluation tool may keep teachers from

embracing the formative attributes. The same document answers the question thus: “The key is to conduct sufficient training and analysis to give *confidence* that concerns with inter-rater reliability will not cause errors in high-stakes employment decisions” (which are associated with the summative aims of evaluation). Again, the concern on developers’ part for teachers’ “confidence” in the rating process and how the evaluation will be used as well as fears regarding summative “high-stakes employment decisions.”

A slide in the *2016 All Faculty Evaluate Davis MicroPD* reminds teachers “Remember, observations are feedback; they are not the evaluation.” This statement seems to be intended to help teachers see that observations are formative and, therefore, to accept the feedback. A column in the *Leadership Training Rubrics* has the heading “Why is it Important?” which implies that the outcomes of the evaluation process are impacted by teachers’ and administrators’ beliefs about the importance of the standards and indicators included in the observation measure.

In the *Evaluate Davis Training for New Educators - Leader Notes* there is additional reference to the impact of believing in the system in the following extract: “We believe the evaluation system provides employees with the opportunity to engage with administrators and colleagues in meaningful conversations about the practice of teaching and learning.” Further data from the same document suggests that the designers of the system saw the need to persuade teachers and administrators of the usefulness of the components of *Evaluate Davis*. With respect to observations:

- Observations allow supervisors to give educators ongoing feedback about their performance throughout the year.
- Observations can give you some good information and, over time, you’ll be

able to see patterns in your observation data, providing you with insight regarding your strengths and areas for improvement.

- Make sure educators *understand* that the observation is NOT the evaluation. The supervisor will use everything they know about the teacher when they complete the end of year summative evaluation. The observations are only a part of that information and provide teachers with feedback on what you actually see or don't see in that 20 minutes of time.

Administrators' interview comments also acknowledged the impact teachers' and principals' attitudes, beliefs, and perceptions have on the efficacy of the evaluation process. For example, Seth asked "...if you're doing it for the purpose of helping teachers..." with the inference that completing the evaluation for the right reasons, with the right attitude, affects the outcome of the evaluation. He shared the case of an administrator going into a teacher's classroom to observe and he's "there for twenty minutes and the teacher isn't even aware" that he's there. He has "some humanity, realizing that everyone has a bad day now and again" and tells the teacher he'll "try again on another day." Seth's case recognizes teachers' plausible poor perceptions of the summative ratings they may receive.

One of Brad's remarks indicated a similar concern for teachers' perceptions of the summative ratings. He explained, with respect to the self-assessment, "What's difficult there is it has a score tied to it. And we don't want to look at the score, or we don't want them (teachers) to focus on the score because it's formative." Seth agreed "...if the teacher is open-minded and is willing to take constructive criticism, as compared to one who thinks 'hey, I know it all and there really isn't room for you to tell me what I'm doing wrong' ...I think if their attitude is 'I'm willing to listen, I'm willing to do that,' then..." they will take the evaluation's formative characteristics seriously. Marshall,

added “If they just say ‘yeah, yeah...I have to do it, I have to listen to the boss telling me what’s going on and get it over with and let me just go back to my room’ ... then nothing changes.” Jill described a thirty-year veteran teacher who responded positively to her feedback. The teacher said, “You know, I really learned something and I realized I wasn’t doing that as much as I should.” Jill said, “That was a cool thing, that he was paying attention and that he was listening and willing to follow through.”

Seth noted that some teachers perceive the whole evaluation process as a perfunctory “hoop to go through” at the end of the year. He said “...if you think of the last evaluation system...there were still goals with that and teachers would write goals you know, but how often did the teachers come back and say ‘Um I’ve got my evaluation, and do you have a copy of my goals?’ At the end of the year, they’re sitting in a file somewhere and...they never looked at them. Well, that still happens now. ‘How do I get back into Evaluate Davis to look at these (goals)?’ Okay, that kind of tells you that the teachers are seeing that as kind of a hoop.” Trent added, “I think too often people don’t take it (goal setting) seriously because they don’t believe in what a goal means... The teachers aren’t bought in and quite honestly, I’m not either.”

Trent explained how principals’ attitudes impact the effectiveness of the evaluation: “How we present it to the faculty, how we follow up, and how much time we take with each of them” affects how teachers respond to the evaluation. Seth added, “I think a lot of that comes back to the attitude we as administrators take towards it...And I think it also goes along with their (teachers’) perception and the experience that they have. If you have an administrator who is just constantly on somebody about changing

and changing and changing or doing something that the teacher constantly feels attacked, then I think that they're going to see that more as... 'let me just do what I do well.' At least in their perception. Again, the teachers that are well functioning, that are open-minded, it could be the way that they were approached with it. So I see it as..." how the "...administrator approaches the individual teachers..." He concluded that teachers' attitude toward evaluation is "...based on our (principals) response to it." Marshall agreed, adding: "I think it does help when you go through it and really think about it as you do"...the evaluation. He said "...it relies on us (principals) to make it so..." teachers see it as formative. Trent added, "...it's, I think, a mentality shift that we (administrators) have to make."

Teachers talked quite openly about their feelings and attitudes about the evaluation process and the impact on how they engage in it. For example, of the formative goal-setting component of *Evaluate Davis*, Karen said "...the last thing I wanted to do is spend time on a goal that I think isn't important." Tom agreed, explaining how the process of goal setting at his school fosters a negative attitude about the process: "...we look at it once a year at the beginning. We all pile into a computer lab, they tell us what to put in and then you don't hear about it again for 365 (he meant 180) days and then all of a sudden everyone's like we've got to get in and do it all again at the busiest times of the year (the end), so it becomes just a complete hurdle that we have to jump over as opposed to a tool to help us get better."

Due to their experience with the previous evaluation system used in the school district, teachers tend to perceive the evaluation system as a way to validate their own

assessment of their teaching and they have become accustomed to receiving the highest possible ratings. They like to hear about the good things observers see, but don't value the ratings they are given. They reported feeling that the ratings were a "reward or punishment," are "threatening," and that not receiving "high scores is disheartening."

As Loraine put it "I love it when the principal comes in and says 'You're doing so good!' It's like yay!" But, she also noted "I don't really care" about the ratings. Nadine expressed her frustration with the ratings as well: "...I don't like the one, two, three, and four scale...the four is supposed to be if you're really good. We're not allowed to have those so it's ridiculous. The district doesn't want evaluators to give them away." She said that with the old evaluation system she received "straight 'you're amazingly awesome'" ratings. Because that is not the case with *Evaluate Davis*, she said the process is "artificial and it negates the validity of any of it... I just think that there's an artificial...hidden agenda in there. So to me I don't care if you give me twos because there's no validity in any of it." The reluctance of administrators to give teachers more than a few "highest scores" (fours) caused Loraine and Tom to adjust their self-assessment ratings: "So what I did to combat it...is when I do my...self-evaluation, I do all fours, like fine, it's all fours. You're not going to give me all fours? I'll give myself all fours," Loraine reported. Tom said he had given himself all fours as well, just to spite his principal. Nadine said that the ratings cause her to dig in her heels. She stated, "Giving me a three or four isn't going to change what I do. No, I'm not going to change my approach for a better score."

Negative perceptions of certain components of *Evaluate Davis* have kept teachers

from embracing the formative characteristics. Nadine expressed her belief that evaluation is a good tool to “identify weaker teachers...to get them out” but, she explained, “If the purpose is really to help teachers, it’s a bigger pain in the butt than I think we get out of it.” Loraine’s distrust of parent and student surveys, likewise, caused her to say they don’t “help me beans.” Nadine added, “It has no meaning.”

Teachers reported that their experience with the goal-setting component of *Evaluate Davis* has not been positive. Tara, for example, said “I honestly have no idea if I’ve even looked at those (her goals) since the beginning of the year.” Karen felt much the same, explaining that her school has teachers gather in a computer lab on a Friday early-out day and input all their goals at the same time: “...we do ours on Friday afternoon at 3 o’clock in the lab together. I’m writing stuff down just so I can leave.” Nadine referred to the goal-setting process as “making up hoops to jump through.” Karen said she sets “easy goals...because I think it’s (evaluation) a hoop.”

Like their administrator counter parts, teachers could see the importance of principals’ attitudes toward the evaluation process in assuring that teachers have positive attitudes with it. Christy explained that the principal’s attitude “can really make a big difference.” She shared her experience with two different principals’ evaluations and described how her most recent principal is “amazing” and has “totally changed the feeling of the whole school.” Tom described a similar feeling with the two very different administrations he has worked with noting that his current evaluations amount to nothing more than “a rating and a ‘good job!’” Karen, too, made a comparison, saying that her new principal “actually went through everything...my feedback was quite valuable this

time, but in other years it has not been.”

Summative aims impact formative aims. As Popham (1988) noted, the summative aspect of evaluation may not be compatible with the formative piece. Many times in the teachers’ focus group interview, teachers complained about the “threatening” nature of the observations and the end-of-year summative evaluations in which teachers are “rated.” “The evaluation lets me know where I stand, but doesn’t help me be a better teacher” is how one teacher indicated her perception that the evaluation is primarily summative. Compounding that feeling is the fear that at some future point, summative evaluations may be used for merit pay or other financial rewards (Pallas, 2010). This fear came out in the interviews when Loraine asked, “Is this heading toward merit pay?” In my own experience, as I introduced Evaluate Davis to my current faculty and staff in 2013, several teachers said they had heard of “school districts in Denver” that were “piloting some form of merit pay” and they wondered how soon that would happen here.

Danielson (2011) reported that educators have become accustomed to seeing the evaluation process as a validation and reward rather than a formative tool for improvement. The teachers I interviewed indicated they were frustrated that they “rarely got a 4 on their observations” and that anything less was “disheartening.” The administrators acknowledged that teachers receiving less-than-perfect scores was “something they need to get used to” if the evaluation is going to be meaningful. As one assistant principal stated “We need to help teachers feel comfortable with evaluation, get used to it. We need to break down the barriers, build relationships, but don’t be afraid to give 2s, either. There is no value to everyone getting high marks.” The current political

climate of accountability exacerbates teachers' feelings that the evaluation is summative rather than formative. As one teacher complained in the interview, "it's the legislature that is doing this to us; they look for any way they can to say we're not doing a good job."

Administrators had a more positive view of the formative nature of *Evaluate Davis*. As one principal put it "I'm having more conversations with teachers than I used to." Another said "I think that just the fact that there's an evaluation tool, it helps with improving teaching...I think it (*Evaluate Davis*) lends to teachers preparing to meet those criteria (34 indicators on which teachers are rated)." One more said "It's more effective in finding what teachers are doing well and where they need improvement...It's beneficial when we work with struggling teachers...we create a plan...we're able to use our evaluation system and observation indicators to say 'these are the areas where you've been weak.'"

Still, there was a tendency among administrators in the focus group to see the summative uses such as rewarding excellence or differentiating between levels of quality of teaching, as these comments suggest: "I think the beauty of this evaluation system is a 3 really is good and a 4 is excellent, and teachers aren't expecting 4s so we're rewarding excellence..." and "...now you can differentiate between those that are really excelling." One administrator noted how the summative evaluation may motivate the teacher to engage the formative nature. He said "I think the summative evaluation is where that conversation can take place about 'Okay, let's do that plan...here are areas that I've seen, or that the administrative team has seen, that you do really well at. Here are some areas

that you could maybe work on' and so then helping the educator develop that professional plan..."

Some administrators expressed a concern about teachers focus on scores rather than the feedback. One said of the summative evaluation "...what's difficult there is it has a score tied to it, and we don't want to look at the score, or we don't want them (teachers) to focus on the score because it's formative." One principal talked about his method for generating summative scores. He said he adds up the scores from all four observations for each indicator and then divides by four and then rounds up. Another principal responded with a more formative approach. He said that he loves "the idea that you're taking the data and using it, but I also look at, instead of just taking an average...if we're kind of going the right way. It doesn't matter when we achieve that mastery as long as we achieve it; that's the score I'm going to give them."

Concern for teachers' attitudes and perceptions of the summative aspects of *Evaluate Davis* are apparent in many of the documents I reviewed. Administrators are trained to assure teachers that their ratings are reliable and fair and are based on more than one or two observations during the year, as indicated from this quote from a training document for administrators:

Discuss the training you've participated in with other administrators to increase inter-rater consistency and any inter-rater reliability work you have done with your administrative team at your school. Most teachers don't know about this and find it both interesting and reassuring to know that administrators are working hard to rate fairly and consistently with other administrators across the district...Make sure educators understand that the observation is NOT the evaluation.

Teachers are also reminded that their supervisors "will use everything they know

about” them “when they complete the end of year summative evaluation. The observations are only a part of that information...” Those who developed *Evaluate Davis* recognized that teachers’ concern over the summative nature of the evaluation would focus their attention on the validity and fairness of the ratings instead of the formative features. For example, administrators are asked to reassure teachers regarding parent and student survey data with this Power Point slide: “Survey data enables us to see ourselves as others see us. Even though that can seem a little scary, it’s good information.” The *Leader Notes* for that slide counsels administrators, “Teachers may ask about the parent response rate.” They are advised to assure teachers that data from parent surveys will only be used in the summative evaluation if an adequate sample of parents complete the surveys.

The rating scale for the summative evaluation has four categories, *ineffective*, *moderately effective*, *effective*, and *highly effective*. Training documents for principals instruct them to help teachers understand “being *Effective* is great! Some people think of *Effective* as a C (grade), but it is not! *Effective* is an A. A *Highly Effective* rating means you are doing something really incredible - just hitting that indicator right out of the ball park!” At the same time, another training document, *Things to Think About: Validity and Reliability*, poses the question for teachers, “If my supervisor is stricter than others, will my *Evaluate Davis* ratings be artificially lowered?” The document answers the question for administrators, counselling them: “The key is to conduct sufficient training and analysis to give confidence that concerns with inter-rater reliability will not cause errors in high-stakes employment decisions.”

Administrators' comments also acknowledged that teachers concerns with summative ratings affect their willingness to accept formative feedback. For example, Seth's comment that when he observes a teacher who is having "a bad day" he tells the teacher he'll "try again on another day" rather than give him low ratings on an observation. Brad also noted a concern for the issues teachers have with summative ratings. He said of teachers' self-assessment ratings: "What's difficult there is it has a score tied to it. And we don't want to look at the score, or we don't want them to focus on the score because it's formative. So I'm regularly having the conversation with them 'Until we have a summative evaluation none of this counts as a score, don't worry about that. If there is a discrepancy between what I'm giving you in an observation and you're self-assessment, that's not a problem.'"

One administrator, Trent, said that although he believes the summative evaluation "has probably given me the opportunity to have more discussions, in-depth discussions, than I ever had" he loves the data that allows him to compare the reliability of his data from observations "with not only me and the district, but me and those that I work with at the school." In this way he can assure teachers that his summative ratings are valid and reliable. Or, as he put it, "when I am working with a remedial teacher and I just need data to back up because they don't understand why we're putting them on this (remediation) plan."

Teachers, too, said that feeling negative about summative ratings affected their willingness to accept and use formative feedback. They used the phrase "I don't care" several times, referring to observation ratings, parent and student survey ratings, and the

end-of-year summative ratings they receive. Words and phrases they used to describe their feelings about summative ratings included “reward or punishment,” “worried about the record,” “dinged,” “hostile,” and “threatening.” They indicated that they fight back with how they complete their self-assessments by giving themselves the highest scores possible without regard for the formative possibilities attached to an honest self-assessment. One teacher, Nadine, went so far as to say she wouldn’t consider changing to get better ratings.

Teachers identified numerous reasons they felt the summative data they received from principals’ observations and parent and student surveys were unreliable: too few parents complete surveys; parent surveys are second-hand - they get their information about the teacher from students; students survey responses are based on how they feel about the teacher at the moment and aren’t really thought out based on the survey questions; observations are twenty-minute snapshots three or four times a year and don’t provide an accurate picture. Overall, teachers don’t like or trust the summative aspect of *Evaluate Davis*. They indicated in their comments that the ratings are superficial and create animosity toward the evaluation process. They do, however, appreciate face-to-face feedback and comments that lead to improvement.

Persistent perception that evaluation is perfunctory. Related to the “summative vs. formative” theme is that of how poorly evaluation is perceived by many educators. For example, during the administrator interview, one principal reminded us that during the first of six half-day administrator trainings on *Evaluate Davis*, we were told that the “most important feature” was how easy it is for principals to use and how

much “quicker to terminate teachers” that need a “change in employment.” Assuring that teachers and administrators use evaluation as an effective formative tool is difficult if their attitude is that evaluation is ineffective, little more than an end-of-year hoop to jump through, or a tool to get rid of poor teachers (Peterson, 2004; Popham, 1998; Tucker & Stronge, 2005). Yet, the tool they are given, *Evaluate Davis*, was presented as just that.

The expectation of receiving the highest possible ratings on observations and end-of-year summative evaluations is the result of many years of experiencing that very scenario. The evaluation process these teachers were most familiar with was one in which 94 percent of teachers receive top ratings (Weisberg et al., 2009). When the system doesn’t differentiate between poor and good teachers, the system is viewed as ineffective, a hoop to jump through at the end of the year (Donaldson, 2009).

Although the documents I reviewed exclaim the formative nature of the evaluation process, other comments within them give the impression that the process is merely a requirement or a yearly task to be completed. For example, school administrators are provided the following script to use when training teachers how to set yearly goals: “Most of the goals fit on a single page and describe elements of best practice instruction and assessment you are most likely already doing.”

The following email to school principals from a district-level supervisor gives the same feeling, especially considering an absence of any written communication advocating for the quality of evaluations:

Reminder-- In looking over the teacher observations you all have performed; I thank you for staying up with these. A couple of schools need to recommit to get these all completed by the end of 3rd term. Oh! That’s today! Principals, would you go over all teacher observations with your team and review who needs to be

observed, knowing that all provisional teachers need 4 minimum observations and then set up a plan to complete these. (Personal communication, March 24, 2017)

Administrators' comments were similar in nature, saying that "sometimes it can feel like" a hoop, and often pointing to the need to "get it done." They expressed the belief that *Evaluate Davis* could be "not just a hoop" if they had "the time...the administrative time to really sit down with teachers and work with them." However, they often find themselves in the situation of running out of time; abandoning efforts to meet with teachers and give quality feedback, they simply aim to get them done. As Dan noted, "I'm not making the time or finding the time to meet with each teacher after" each observation. He said, "We're very busy. If every kid was perfectly behaved I think we'd have time to sit down and meet with those teachers." He added that at the end of the year he's "got 6 pages taped to the back of my door and I've got color coordinated marks for the terms they were observed and we're checking them off. I love checking them off, so that's a hoop."

Seth noted that at the end of the year when he meets with teachers, the evaluation sometimes "feels like a hoop." He talked about teachers who set goals at the beginning of the year, put them "in a file somewhere and...they never looked at them." At the end of the year they ask "How do I get back into *Evaluate Davis* to look at them?" At the end of the year is not the only time when the goal setting piece feels perfunctory. Trent pointed out that for teachers, "their curriculum advisor has set" their goals for them, and they haven't "changed since last year" so they are "exactly the same and there's not a lot of value in those goals." Dan expressed a similar experience, describing teachers' goal-setting as a "cut and paste" process. He said, sarcastically, "Whoever is writing them at

the district is doing better than my teachers.”

Teachers admitted that to them the evaluation process often feels like a once a year hoop to go through. Karen said, “I think it’s a hoop. I do.” Tom’s comment summarizes teachers’ comments from the interview: “...we look at it once a year at the beginning. We all pile into a computer lab, they tell us what to put in and then you don’t hear about it again for 365 (he meant 180) days and then all of a sudden everyone’s like we’ve got to get in and do it all again at the busiest times of the year (the end), so it becomes just a complete hurdle that we have to jump over as opposed to a tool to help us get better.” Nadine described *Evaluate Davis* as a “bigger pain in the butt...a much bigger hoop to jump through and pain...”

I asked teachers if the goal setting process felt like a hoop. They answered emphatically “Yes!” Tara added “The goals for sure. I honestly have no idea if I’ve even looked at those since the beginning of the year.” Karen told of how her school has teachers gather in a computer lab on a Friday early-out day and input all their goals at the same time: “...we do ours on Friday afternoon at 3 o’clock in the lab together. I’m writing stuff down just so I can leave.” Teachers complained that the goal setting component of *Evaluate Davis* is a “hoop to jump through” because “...we get goals given to us.”

As one of the six teachers in the focus group reported, “I can play the numbers game, do all the things to get a score, but why? I don’t care. They’re only in my room for twenty minutes or so; they don’t see most of what goes on and then I don’t get a chance to tell them what they missed.” Focused on the score or rating, this teacher missed the

point of formative evaluation altogether - to assist teachers in becoming better at their job of getting all students to learn. Marzano et al. (2011) made that very point that teachers need to see and accept the evaluation process as formative and be willing to receive less-than-perfect scores as a way of identifying specific areas of needed improvement.

The data from the interviews and document review, and what I learned from the literature review point to a number of factors which influence the persistence of the negative perceptions teachers and administrators hold of teacher evaluation. *Evaluate Davis*, and other similar evaluation tools, still do not take into account principals lack of time and training. As a result, most teachers still receive good or excellent ratings, indicating the instruments are less than effective at differentiating between effective and ineffective teachers. Training materials still focus on the mechanics of the evaluation process, such as how many observations to do and how quickly feedback is given, rather than the quality of the evaluative experience, such as the value of the feedback given. Training materials still imply that teachers are evaluated to comply with legislative mandates. And, important formative procedures, such as self-assessment, goal setting, and professional development are rushed through and aren't required to align. Principals are not trained to facilitate the process and don't take the time to work with teachers to assure alignment, and teachers don't see the need for alignment since it isn't required or focused on in training materials.

On a larger scale, another huge factor is that the evaluation process is a legislative mandate which feels top-down. Federal legislation has had a big impact on teacher evaluation—46 states have revamped their evaluation systems as a result—and the fact

that evaluation is required by law impacts teachers' and administrators' perceptions and attitudes. Indeed, as I have noted previously, school administrators in Utah are required by state code to evaluate their teachers yearly using the newly accepted, more rigorous standards required by Federal initiatives, such as Race to the Top and the ESEA Flexibility Program (Steinberg & Sartain, 2015). Although Davis School District followed recommendations (Goe, 2013) to heavily involve teachers and school administrators in the development of *Evaluate Davis* in order decrease the top-down feel of legislative mandates, negative perceptions of evaluation still persist. Teachers said more than once in the interviews that we only have *Evaluate Davis* because of the legislature. Furthermore, the final product still needed approval from the Utah State Board of Education to assure it met the standards required by Federal initiatives. Some of those very standards are what teachers complained about most in the interviews: parent and student surveys, measures of student achievement based on standardized tests, and scores for observation items.

In the following chapter, I summarize the conclusions of the data analysis. Data that were gathered through interviews, document review, and the literature review, are presented in the context of each research question. Findings and themes are discussed in terms of the overall aim of this study and how they relate to current literature. Recommendations for improving the likelihood that *Evaluate Davis*, and comparable evaluation systems, will achieve their objective of improved teaching and increased student learning, are presented, as well as recommendations for further research.

CHAPTER V

CONCLUSIONS

Findings

My purpose in conducting this study was to construct a description of teachers' and principals' attitudes, beliefs, and perceptions regarding a teacher evaluation tool, *Evaluate Davis*, and gain insight into how their attitudes and beliefs impact the implementation of the evaluation. Through this study, I sought to gain an understanding of how teachers' and administrators' perceptions of *Evaluate Davis* influence the evaluation process. I explored Teachers' and administrators' attitudes, beliefs, and perceptions of *Evaluate Davis* using focus group interviews and document reviews. Data were analyzed and predominate themes exposed to assess the influence of attitudes and perceptions on the implementation of *Evaluate Davis*. Results of the analysis may be used to create an optimal approach for implementation of *Evaluate Davis* and similar evaluation tools with the aim of increasing the likelihood of teacher growth and improving teaching, increasing student learning, and improving schools. Related questions that guided this study were as follows.

- What are administrators' and teachers' perceptions of the role *Evaluate Davis* plays in improving student learning gains?
- What are the perceptions of teachers and administrators respecting how *Evaluate Davis* is used to improve teaching and learning?
- To what extent do teacher and administrator attitudes regarding the efficacy of *Evaluate Davis* impact its usefulness in improving teaching and learning?

Discussion of Results

In this section, I summarize the conclusions of the data analysis. Data that was gathered through interviews, document review, and the literature review, are presented in the context of each research question. Findings and themes are discussed in terms of the overall aim of this study, the theoretical framework of the study, and how they relate to current literature.

The data analysis led to several conclusions.

1. The negative perceptions of the teacher evaluation system employed previous to the implementation of *Evaluate Davis* have persisted, even though the systems are very different. This appears to be the case for a number of reasons. One is that even though the mechanics differ, it is still implemented the same way as the previous evaluation model by most principals and district officials. Principals still don't have the time or training they need to be effective evaluators. They still give ratings of good or excellent to nearly ninety percent of teachers. Training materials still imply that the process is in place to comply with state mandates. And, procedures such as self-assessment, goal setting, and professional development plans are not required to align.
2. The expectation of policy makers for *Evaluate Davis* to become a key factor in improving teaching in the school district and increase student learning is quite different from how teachers and principals view *Evaluate Davis*. Principals see the potential for the evaluation process to help teachers improve, but feel frustrated by the many barriers they see - chiefly lack of time and training. Teachers perceive the evaluation as primarily summative, and even punitive, and expressed the view that little about it motivates them to improve.
3. The summative role of *Evaluate Davis* is a barrier to teachers' acceptance of critical feedback. Evaluation needs to be focused on teacher improvement and growth.
4. The principal's attitude, beliefs, and perceptions are critical to the effectiveness of the evaluation process; how they implement it is how teachers will receive it.
5. Principals' ability to develop relationships of trust with teachers is also a key variable in the success of the evaluation process.

6. The evaluation process with the greatest potential for success involves principals being in teachers classrooms often, both formally and informally, followed up with face-to-face feedback, not only about what they saw, but also with suggestions and solutions; teachers value instruction-focused dialogue. Furthermore, dialogue should lead to administrators working collaboratively with individual teachers to develop improvement plans and professional development.

First research question. In the first research question, administrators' and teachers' perceptions, along with beliefs and perceptions evident in district documents, were explored as they pertain to the role *Evaluate Davis* plays in improving student learning gains. Data from all three sources - documents, teacher interviews, and administrator interviews - reflect differing views of the role evaluation plays in improving teaching and increasing student learning. Even within each data source there was some disagreement. For example, principals stated they believe the most important characteristic of *Evaluate Davis* is the tool's ability to differentiate between good and poor teachers, but also believed evaluation was making a difference in improving teaching in their buildings. However, district documents were unambiguous in proclaiming that the most important aim of the evaluation process is teacher growth, with the belief that increased student learning will result. Teachers, like the principals, felt the role of *Evaluate Davis* was primarily to "identify poor teachers" but they also said that most teachers would not improve their teaching because of evaluation.

The documents' focus on teacher improvement and increased student learning as the role of evaluation is congruent with references found in the literature review that recommend the goal of evaluation should be on student learning through development of teachers' instructional ability (Darling-Hammond et al., 2012; Donaldson, 2009; Goe et al., 2008; Havens & Roy, 2007; Hazi & Rucinski, 2009; Towe, 2012). School

administrators in the interview were balanced in their perceptions, seeing the role of evaluation as both summative and formative, which is also supported in the literature review. For example, Darling-Hammond et al., (2012) concluded that the best teacher evaluation systems are focused on helping teachers improve, but also support timely and efficient personnel decisions. In addition to differentiating the quality of teachers, the principals indicated they could see great potential for evaluation to help improve teaching in their schools and believe that in doing so, student learning will improve. Teachers' perceptions were less optimistic and reflected the findings of current research showing a lack of evidence that evaluation reforms are improving teacher performance (Darling-Hammond et al., 2012; Kraft & Gilmour, 2016). Differing perceptions of the role of the evaluation process held by those who prepared the training documents, those who administer the evaluation in schools, and those who are evaluated, presents a potential barrier to the successful implementation of *Evaluate Davis*.

Second research question. In the second research question, administrators' and teachers' perceptions of how *Evaluate Davis* is used or implemented were explored. Halverson et al. (2004) completed a case study in which teachers and principals were interviewed to identify their perceptions and attitudes toward the evaluation process. A range of perceptions came out of the interviews, from evaluation is "an opportunity to develop morale" to evaluation is nothing more than a "mandate" that creates a time-management problem for principals. They found that both teachers and administrators viewed traditional evaluations as a perfunctory burden, used in most instances to get rid of teachers with the poorest performance. The evaluation process, they found, did little to

encourage teachers to improve and few teachers claimed to have made any changes in instructional practice as a result of evaluation. The study reported widespread belief among teachers and administrators that evaluation is “not a primary force for improving teaching” (p. 12).

My experience interviewing teachers and administrators uncovered a similar variety of perceptions. The teachers and administrators I interviewed saw evaluation as primarily for identifying poor teachers, but also felt it should be used to validate and reward good teachers. They agreed that it often feels like a yearly hoop to jump through, a mandate from the state that is both time and resource consuming. However, while teachers were emphatic that the evaluation process did nothing to improve their teaching, administrators were just as insistent that it is making a difference in teaching and learning.

The district training documents I reviewed were very specific in the actual strategies, procedures, and practices school leaders and teachers are expected to follow in the evaluation process. However, the actual implementation of those strategies, procedures, and practices was very different from school to school depending on the school administrators’ attitudes and beliefs about the evaluation process. Both teachers and administrators acknowledged that the outcome of evaluation depended on the perception of the principal.

For example, among the administrators who participated in this study, those who viewed *Evaluate Davis* as primarily summative were less likely to take the time to give constructive feedback or engage teachers with face-to-face conversations about their

teaching. More often, they gave teachers a rating and relied on the teacher to seek them out if they disagreed with it. On the other hand, those administrators who see the process as formative reported giving more detailed comments directly on the observation protocol and said they were more likely to seek out opportunities to meet directly with individual teachers to give feedback and engage in constructive instructional dialogue.

Teachers' responses verified an array of administrators' evaluation practices ranging from administrators who simply tell teachers they know they are good and don't need to be observed, to those who complete evaluations just to get them done, to those who actually make comments on the protocol and give beneficial feedback. Teachers reported their favorite part of evaluation is having administrators come into their classrooms to see what's going on there and then giving face-to-face feedback that includes positive observations as well as solutions and suggestions for improvement. They like the comments they get on the observation protocol and feel most inclined to respond to those, but also felt that in some cases, comments and feedback from their administrators were lacking. Teachers said that some principals did nothing more than give ratings on the checklist, which they resented. Teachers viewed the observation rating scale as mostly summative and disliked the summative evaluation ratings, as well. They acknowledged becoming accustomed to receiving high ratings on the checklists and that low scores "just make them mad." They also expressed frustration that administrators are encouraged by district supervisors to limit the number of high scores they give teachers.

Not only did teachers not trust the observation rating scale, they expressed concerns about all of the other components of *Evaluate Davis*. Teachers reported they

don't like or trust the summative aspect of *Evaluate Davis*, calling it "threatening" and "hostile," a "hoop to jump through" at the end of the year. They indicated in their comments that the ratings are superficial and create animosity toward the evaluation process. Teachers did not express much confidence in the other forms of data that are used for summative evaluations, despite the assurances present in district documents to persuade them otherwise. They talked of numerous concerns with parent and student surveys, classroom observations, and the self-assessment and goal-setting process. They didn't feel that parent and student surveys were used formatively, sharing concerns about the fairness and validity of survey data, especially when used for summative purposes. They also complained that the goal setting component of *Evaluate Davis* is nothing more than a "hoop to jump through." They do, however, appreciate face-to-face feedback and comments that lead to improvement, despite saying that evaluation did not help them improve.

The teachers' comments agreed with what was learned from the literature review regarding the summative aspects of evaluation impacting the formative. Popham (1988) argued that summative and formative purposes of evaluation are not compatible, claiming that summative uses undermine the formative. Gleave (1997) reported that attempting to use the same evaluation system to measure teachers' effectiveness and to help them improve only creates frustration and resentment, as Hazi and Rucinski (2009) noted, often exacerbating poor relationships between teachers and administrators. Popham (2013) described the impossible task of being both a summative and formative evaluator that school administrators are faced with. He provided evidence that when evaluation

takes on a summative role, the formative role is weakened. He noted further, that when teachers feel threatened by summative consequences they are less likely to accept formative feedback in a positive way. Additionally, there is general agreement among researchers and educators, according to Darling-Hammond et al. (2012), “that most current teacher evaluation systems do little to help teachers improve or to support personnel decision-making” (p. 8).

Teachers and administrators both talked about the need to develop relationships of trust between teachers and administrators. Both focus groups reported the need to find the time to engage in meaningful dialogue regarding instructional practice and to collaboratively identify areas of strength and weakness and develop a subsequent plan for improvement. Goe (2013) expressed the need for principals to become experts in maintaining positive relationships with teachers and in helping them monitor and evaluate their own growth and practices. Similarly, Towe (2012) advocated for principals to maintain positive relationships with teachers, explaining that they need to help teachers see evaluation as useful and fair and assure that evaluations take place in an environment of support, collaboration, reflection, and professional development.

Principals in the focus group acknowledged they could benefit from additional, ongoing training to help them become more proficient and consistent evaluators, and in developing and maintaining positive relationships with teachers. They expressed frustration throughout the interview with perceived barriers, such as a lack of time and training, which left them less optimistic about improved teaching and student learning. Principals reported they sometimes give little or no meaningful feedback, waiting instead

for teachers to come to them if they have questions or concerns. They noted that a lack of time kept them from engaging in meaningful conversations with teachers, completing more frequent classroom observations, and completing the observation protocol effectively.

Teachers also felt too little time was spent on providing meaningful feedback. They reported that, to save time, principals' feedback is often too general and lacks the specificity to be helpful in improving instruction. They also expressed concerns with inter-rater reliability, reporting that the ratings they receive often depended on who was evaluating them. Some administrators and teachers saw the actual evaluation process as cumbersome and unwieldy. In both focus group interviews, principals and teachers reported that principals resort to checklist observations to save time and "get observations done." Nonetheless, principals were optimistic that *Evaluate Davis* was making a difference in teaching and learning in their schools.

The current research also recognizes comparable barriers. Time is a big issue for principals, according to recent literature, especially in secondary schools. In Towe's (2012) research, principals reported a lack of time to observe teachers, for conferencing with teachers, and to provide quality feedback. According to Halverson et al. (2004) principals put a lot of effort into implementing all of the components of the evaluation, but find it time-consuming. Principals, according to Donaldson (2010), have traditionally been poorly trained in evaluating teachers, and explained that a lack of direction from districts, including clear guidelines and rubrics for principals, contributes to poor quality evaluations.

Halverson et al. (2004) found that principals rarely get feedback from their supervisors about whether they are doing a good job as an evaluator, even though relevant experience and training are seen as important to evaluators' success at assessing teachers. A lack of training results in principals being less than skilled in their ability to implement components of the evaluation, especially the critical elements of gathering and interpreting data and following it up with quality feedback (Goe, 2013). Steinberg and Sartain (2015) found that the efficacy of evaluation to improve teaching depended in part on the principal's competence in providing directed instructional assistance. Halverson et al. (2004) suggested that evaluators and teachers need greater amounts of training on how to collect data, reflect on it, and present it in ways that increase its potential for improving teaching and learning. Additionally, they recommend training focused on helping evaluators develop their skill and will for giving critical feedback.

Donaldson (2010) explained that a lack of direction from districts, including clear guidelines and rubrics for principals, contributes to poor quality evaluations. I found that inconsistencies in the focus of the training materials provided by the school district may lead both teachers and school administrators to be confused about how *Evaluate Davis* is to be used. Although much of the training found in the documents is focused on the formative, there is also a clear concern for teachers' anxiety over the summative facets of the evaluation. For example, many of the *Leader Notes* principals are encouraged to use when training teachers about the evaluation process are focused on helping teachers feel comfortable with and confident in the summative ratings they receive from evaluators, parents, and students. On the other hand, there is little mention in the documents about

how principals can give meaningful feedback or engage teachers in instructional conversations that lead to formative outcomes. Furthermore, some of the data found in the documents, such as references to state mandates, serve to reinforce perceptions that the evaluation process is a hoop to jump through to be in compliance with policy regulations. Hazi and Rucinski (2009) stated that state mandates regarding evaluation policies and procedures will reinforce the reputation that teacher evaluation is a meaningless ritual.

I found another interesting contradiction that adds to the perception that the evaluation process is just a yearly hoop to jump through. Although the training materials advocate a pattern of teacher improvement spawned by the evaluation process that includes teachers and administrators gathering data through observations, parent and student surveys, and self-evaluation, then collaboratively setting goals based on that data and developing a personal professional development plan tailored to their goals, there is also a slide in the same training materials indicating that teachers may use pretty much anything they want as professional development: School-based PD, district-based PD, graduate coursework, independent study, book study, etc. There is no mention that it needs to be linked to their goals. So, teachers reported they use whatever they are already doing, irrespective of whether it will help them improve in identified areas.

At best, such seemingly contradictory messages foster interesting differences in principals' approaches to the evaluation. For example, one principal arrived at summative ratings for his teachers by averaging all of their scores from observations and parent and student surveys. Another principal took a more standards-based approach. He looked at

the latest ratings teachers received and if they showed improvement from the first to the last, he used the last scores given. Although principals seemed comfortable that *Evaluate Davis* could meet both summative and formative aims effectively, teachers perceived it as largely summative and didn't believe it was serving both purposes effectively.

Summarily, notwithstanding district training materials that emphasize the formative aims of *Evaluate Davis*, and administrators' confidence that teaching will improve as a result of the evaluation process, teachers were much less enthusiastic. They expressed skepticism that any of the components of *Evaluate Davis* could make a difference in their teaching or student learning. They believed that although the proclaimed purpose is formative, they see it as largely summative. The summative role they said prevented them from fully engaging in the formative aspects.

Third research question. Regarding research question three, data from the document review indicate that those who designed *Evaluate Davis* recognized that teachers' attitudes and perceptions of the evaluation's efficacy is important with respect to the expected impact it may have on teacher improvement. Historically, teachers' views of evaluation have been somewhat negative (Pallas, 2010; Peterson, 2004). Halverson et al. (2004) found that both teachers and administrators viewed traditional evaluations as a perfunctory burden, used mostly to get rid of teachers with the poorest performance. Current literature advises that attitudes toward present evaluation procedures are similar; many teachers, principals, and researchers are still suspicious of recent changes to teacher evaluation systems comparable to *Evaluate Davis* and are unsure of their value in terms of improving teaching (Goodwin & Hein, 2016; Kraft & Gilmour, 2016; Popham, 2013).

The teachers I interviewed expressed equally negative perceptions of *Evaluate Davis*.

Halverson et al. (2004) assert that the perceptions held by teachers and administrators of the evaluation system determines how successful the evaluation procedures will be in affecting improvement in teaching and learning. Teachers' attitudes, likewise, influence how they receive the evaluation process and react to feedback (Kimball & Milanowski, 2009). According to Halverson et al. poor perceptions of evaluation have had a negative effect on efforts to use evaluation to impact student learning in positive ways. They asserted, as well, that the perceptions held by teachers and administrators of the evaluation system determine how successful the evaluation procedures will be in affecting improvement in teaching and learning. Likewise, Marchiando (2013) found that the perceptions of teachers and principals toward the evaluation process have a direct impact on the system's implementation and effectiveness "as a tool for growth, which ultimately impacts student achievement" (p. 4). Fundamentally, current literature is clear that educators' attitudes about evaluation impact, for good or bad, how the evaluation is implemented.

The training documents reviewed in this study encouraged teachers to "care about their performance" and to "strive" to improve. Many of the training documents were focused on training administrators to persuade teachers why evaluation is important, the usefulness of each of the components of the evaluation process, and to "believe the evaluation system" provides teachers with opportunities and tools to improve "the practice of teaching and learning." Teachers are reminded in the training to "Take feedback from your supervisor seriously, and work to implement their suggestions for

improvement.”

The teachers and administrators who participated in the focus groups agreed that their own attitudes and perceptions influenced their engagement in *Evaluate Davis*. Both teacher and administrator interviews confirmed that administrators’ perceptions, attitudes, and beliefs influenced the way teachers perceived the evaluation process and how it is received in schools. Principals and teachers acknowledged that teachers’ attitudes about the evaluation are largely dependent on their principal’s attitudes and approach. Both groups said that if principals viewed evaluation as formative, then they took more time helping teachers focus on growth and improvement and they spent more time in giving quality face-to-face feedback and in engaging in conversations about quality teaching with teachers. Teachers, in turn, said they were more likely to respond positively to that approach. Both groups agreed that when principals seemed more interested in developing trusting relationships with teachers, teachers were more apt to grow professionally from the evaluation process.

Theoretical Framework

In Chapter II, I discussed the theoretical framework that guided my study. I explained how Moore’s *Theory of Transactional Distance*, Roger’s *Diffusion of Innovations* theory, and Marzano’s *Evaluation Model* directed my research. This same framework informed the findings and conclusions of my study, as well. Moore’s *Theory of Transactional Distance* describes the “space of potential misunderstanding” and “lack of mutual understanding or common perception of ideas, emotions...knowledge, approaches” (Giossos et al., 2009, p. 3) which define the gap, or distance, between how

teachers and administrators, as the recipients of the evaluation program, experience the evaluation process, and the expectations of policy makers and evaluation designers. The findings of this study indicate that the gap of misunderstanding still exists. While policy makers and those who designed *Evaluate Davis* clearly believe, as the Marzano *Evaluation Model* describes, that the evaluation process has the potential to help teachers to incrementally improve in skill and expertise, teachers and principals still view it largely as a perfunctory exercise. On another level of the process, there is a transactional distance between teachers and evaluators that “affects behaviors in major ways.” At both levels, the degree of separation effects the delivery of the evaluation and how it is received by teachers. The kind of dialogue between teachers and evaluators is evidently determined by the educational philosophy of the evaluation designers who created the evaluation system and is influenced by the personalities of teachers and evaluators as well as environmental factors in the individual work places. The structure, or rigidity of the evaluation tool’s objectives, implementation strategies, and methods clearly impact the degree to which the evaluation system can be responsive to individual teacher’s needs.

Roger’s *Diffusion of Innovations* theory was foundational to my study on two fronts: First, the history of teacher evaluation follows the pattern described above by Valente and Rogers (1995). As the literature review demonstrates, evaluation systems have changed over time as educational paradigms have changed. And second, in this study I examined a teacher evaluation system, *Evaluate Davis*, which is a variety of the new evaluation systems being implemented throughout the U.S. The new systems are

considered to be innovations in and of themselves, but are also based on the paradigm, described in Marzano's *Evaluation Model*, of teacher growth as a means of improving student learning. As I sought answers to my research questions, I examined the perceptual distances present in the widespread "diffusion" of a new evaluation process that includes a paradigm of evaluating teachers based on a model of teacher improvement.

Greenhalgh et al. (2004) defined diffusion of innovations as it pertains to health service organizations, but the definition is applicable to educational organizations also, as paraphrased: "...a novel set of behaviors, routines, and ways of working that are directed at improving *educational* outcomes, administrative efficiency...or users' experience and that are implemented by planned and coordinated actions." They explain that implementation refers to "...planned efforts to mainstream an innovation within an organization" and sustainability refers to "making an innovation routine until it reaches obsolescence" (p. 582). In the recommendations proposed in Chapter V, I advocate for improved methods of implementation that may increase the sustainability and formative success of the evaluation process.

As the theory of *Diffusion of Innovations* applies to the implementation of a new teacher evaluation system based on a new paradigm, the approach for spreading the innovation in a school district is largely active dissemination. Specifically, in the case of *Evaluate Davis*, certain attributes were mandated by the state legislature to assure that the evaluation would be in compliance with the most recent recommendations of researchers and experts. Davis district-level administrators carefully planned and carried out the

actual implementation of the system at the district level and have worked to assure that it is implemented with great fidelity (the same in each school) in every school in the district. However, the implementation at the school level is somewhat haphazard, as some principals have put a lot of time and energy into training their faculty and preparing them for the more rigorous evaluation system, while others have taken the pure diffusion approach with little planning or training.

The *Marzano Evaluation Model* is another important part of the theoretical framework that supported my study. Marzano's model is built on the foundation of iterative improvements to the evaluation and supervision of teachers over the history of public education in the U.S. and on a strong research base supporting the following three points: (a) Teachers who are more skilled and expert produce greater gains in student achievement; (b) The evaluation process has the potential to help teachers to incrementally improve in skill and expertise; and (c) The methods for improving teacher skill and expertise are clear (Marzano et al., 2011). It is established on the principle that "the purpose of supervision should be the enhancement of teachers' pedagogical skills, with the ultimate goal of enhancing student achievement" (Marzano et al., 2011, p. 2).

Summarily, the Marzano model of teacher evaluation is based on the principle that "evaluation is to enhance teacher effectiveness." In practice, the model relies on multiple forms of feedback as the foundation for encouraging teacher growth in specific teacher skills. The system uses multiple measures of teacher growth in a "rigorous and informative approach" (Marzano et al., 2011, p. 103) that is intended to motivate professional development and recognize the different stages of progression toward

expertise.

Summary of Findings

Although *Evaluate Davis* was designed to fulfill both summative and formative purposes, the greater aim is formative. This is clear in the documents I reviewed and in the views and opinions expressed in the focus group interviews. Despite clearly stated goals of teacher growth and improved teaching, some implementation strategies and training tools can easily lead to the conclusion that the goal of evaluation is to comply with state laws. Many of the practices and procedures school leaders engage to complete the evaluation process appear aimed at getting the evaluation done, as well.

Consequently, long-held perceptions of teachers and administrators, that the evaluation is nothing more than a yearly hoop to jump through, continue to persist.

The perception that the evaluation process is a perfunctory measure impacts the evaluation process at its most critical level, where administrators give feedback to teachers about their observation of classroom teaching. How that feedback is given is dependent on the administrator's perception and attitude about the evaluation process. Likewise, the teacher's reception of the feedback is dependent upon his or her attitude and beliefs. Principals' and teachers' attitudes, perceptions, and beliefs about evaluation are impacted by a number of factors: Time, principals' will, principals' skill and training, how the process is presented and used (formatively or summatively), teachers' training with the evaluation tool, consistency or inter-rater reliability, and differing opinions of what quality teaching is.

Teachers' view of evaluation, in general, is negative. The very component they

value and trust the most - face-to-face feedback with administrators and conversations about their teaching - is the most lacking in the process, due to the constraints principals identified: lack of training and lack of time. Teachers complained about each of the components of *Evaluate Davis*. They don't trust parent and student surveys; they believe observations are too infrequent and short; they admitted their self-assessments are often contrived and used to make up for perceived-poor observation ratings; and they reported that goal setting has been marginalized by being taken over by district curriculum supervisors.

Teachers' perception of the evaluation process as mainly summative negatively influences their desire to engage in the process. Teachers' lack of trust in their administrators' assessment of their teaching also hinders their willingness to accept feedback and make positive changes that improve their teaching. In general, the expectations district and state policy makers have of the evaluation process and how it impacts teaching and learning are quite different from how school level administrators and teachers are experiencing it. District level documents indicate a strong, clear focus on teacher growth and improved teaching as the aim of *Evaluate Davis*. School leaders desire to realize the formative aim of evaluation and recognize the possibilities, but feel the very real constraints of time, lack of training and skill, and difficulty in providing consistency.

Recommendations

Teachers, according to the literature review, are the most important factor in

assuring that all students learn and achieve school success (Akiba et al., 2007). Assuring that every student is taught by highly effective teachers is a noble goal and seems to be the focus of the decade-old trend to revamp teacher evaluation systems throughout the U.S. (Donaldson & Papay, 2015; Goe et al., 2008; Kimball & Milanowski, 2009; Popham, 2013; Steinberg & Sartain, 2015). Teacher evaluation systems have evolved into two-pronged arrows aimed at improving teaching in America's schools, with the assumption that student-learning gains will increase. One prong uses formative efforts to engage teachers in self-improvement. The other prong uses summative processes to make employment decisions that remove poor teachers from our schools (Popham, 2013).

Evaluate Davis is a teacher evaluation system recently implemented in Davis School District, a large suburban school district in Utah. Although the stated intent of *Evaluate Davis* is to improve teaching in the district, in focus group interviews, teachers and administrators acknowledged a number of perceived barriers to accomplishing that aim. The following recommendations for overcoming those barriers and improving the likelihood that *Evaluate Davis*, and comparable evaluation systems, will achieve the objective of improved teaching and increased student learning, are based on the literature review, the document review, and the teacher and administrator interviews.

Perception Replacement

The perceptions, beliefs, and attitudes of educators are critical to the effectiveness of the teacher evaluation process. Principals' attitudes about evaluation impact, for good or bad, how they carry out their evaluation duties. This point was clear in both the principal and teacher interviews. Teachers' attitudes, likewise, influence how they

receive the evaluation process and react to feedback (Kimball & Milanowski, 2009).

Culture. The collection of the beliefs, attitudes, and values of the educators in a school or district is described by DuFour (1998) as the culture of the school or district. A premise of this study is that when a school or district seeks to implement any large-scale change, such as a new evaluation system, the culture must be addressed (National Association of Secondary School Principals [NASSP], 2006). The actions, traditions, symbols, ceremonies, and rituals of the district must be addressed to assure they are creating the culture conducive to the aims of the change to be implemented; in this case, a new evaluation system (Fisher, Frey, & Pumpian, 2012).

District and school leaders must work to create a change in culture, from the one that views evaluation as an end-of-year, government sponsored, perfunctory exercise, to a culture that believes that the evaluation process is a valuable formative tool and important to teacher improvement. District leadership and training should focus on bringing school leaders and teachers to see the formative value in the evaluation process. For example, all of the components of *Evaluate Davis* that are recommended to assure that the evaluation actually promotes teacher improvement and student learning gains can be implemented with fidelity, and yet the evaluation will fail to accomplish that goal if both the teacher being evaluated and the administrator doing the evaluating don't believe in the goal and the importance of the process in achieving the goal.

Articulate, demonstrate, and nurture. As noted in the interviews and the literature review (Darling-Hammond et al., 2012), the most important aim of evaluation should be teacher growth and increased student learning. Principals should receive training to learn

to articulate and demonstrate strong belief in the goal of teacher improvement and to encourage teachers' confidence in the formative benefits of the evaluation process. They must learn to nurture teachers' commitment to engaging in the process of improvement with high expectations of a positive outcome and to welcoming the formative possibilities built in to the evaluation process (DuFour, 1998). The ideal of assuring that every student has access to highly effective teachers should be the focal point of evaluation and all involved need to believe that the outcome of that process is more students learning more, more college-ready kids, higher graduation rates, and fewer failures.

Get on the same page. During the interviews, some principals expressed some confusion about the purpose of Evaluate Davis, while teachers were certain the aim is mostly summative. The training documents are clear that the aim is teacher growth, however. School district leaders must ensure congruence within all stakeholder groups (policy makers, school administrators, and teachers) and documents that the purpose of evaluation is teacher growth and improvement aimed at increased student learning gains (Havens & Roy, 2007; Marzano & Toth, 2013). References in documents and training materials that lead to the perception that evaluation is something teachers and administrators are forced to do by policy regulation should be extracted. Communications regarding teacher evaluation, between stakeholder groups and within stakeholder groups, should be carefully crafted to avoid reinforcing negative perceptions of teacher evaluation and, rather, to perpetuate formative goals of growth and improvement.

Much of the time principals spend in training is spent on the mechanics of implementing *Evaluate Davis*. Principals learn about how many observations they need

to complete, how many times they need to have conversations with teachers, how many goals teachers need to set, and so forth, but never get to the meatier matters of what quality teaching looks like. More time ought to be given to training teachers and administrators regarding the district-provided rubrics for the observation indicators. These rubrics provide a framework for understanding effective teaching; they foster a common language, identify specific teaching strategies and behaviors, and clarify expectations of teacher growth and improvement. They serve to eliminate confusion regarding what specific observation indicators mean and help to assure greater understanding between evaluators and teachers. The rubrics allow for more specific and timely feedback tied to research-based teaching practice, as well as better self-assessment and reflection centered on the development of specific improvement goals and individualized professional development plans.

Practices and policies. Practices and policies, which reinforce the perception that any evaluation components are a formality to get out of the way, should be replaced with authentic procedures that teachers and administrators believe have formative value. For example, to streamline the goal-setting process for teachers, school district curriculum supervisors have taken on the task of providing goals for teachers, rendering the goals, according to the teachers interviewed in this study, meaningless to them. Teachers in the study recommended that teachers' goals need to be personal to them, based on an examination of data from administrators' observations and feedback along with their own perceptions of their teaching effectiveness. Goals, they said, ought to match identified areas of needed improvement and fit within a broader framework of school, department,

and team aims, and should lead to the crafting of a plan for professional development.

Self-efficacy. Administrators would do well to focus on building each teacher's sense of self-efficacy (belief in their ability to impact student learning in positive ways). Marzano et al (2011) reported that the effectiveness of teacher evaluation systems rests on teachers' acceptance of the process as formative and their willingness to accept less than perfect scores as a means of identifying areas of needed improvement. One of the attributes of teachers with a strong sense of self-efficacy is a willingness to accept feedback that encourages more experimentation with new ideas and methods to meet more students' needs (Prothero, 2008). Bruce (2008) noted the important relationship between teacher quality and teacher self-efficacy; "Teachers" she said "must focus on their own attitudes and classroom behaviors..." (p. 1). Teachers' ability to impact student learning in positive ways is affected by their own beliefs about their potential to influence student learning (Goddard, Hoy, & Woolfolk, 2000; Torff, 2011). Principals should receive training on how to build teachers' capacity and self-efficacy as a prerequisite to expecting them to accept evaluation as formative.

School administrators must be taught to cultivate and nurture the seeds of formative evaluation. They should focus their discussions with teachers and their evaluative efforts on the formative aspects. They should receive the training they need to help them find time to meet face-to-face with teachers, give quality feedback that includes rich discussions about enhancement of teaching strategies, and develop their own expertise in pedagogy. They need to learn how to gain teachers' trust and confidence in their ability as expert evaluators. District leaders, school leaders, and teachers must

acknowledge that effective teachers are the most important factor in student achievement and that therefore, by improving the effectiveness of teachers, student learning will increase. Furthermore, educators must be convinced that the evaluation process is a crucial component of the teacher improvement process. All involved in the process must see the value of evaluation and be willing to both give and receive feedback with an attitude of continual change and growth.

Principal Knowledge and Skill

The principal's leadership is crucial in the effective implementation of the evaluation process. The burden of teacher evaluation rests primarily with principals who determine the usefulness of the system through their implementation of the system's components (Steinberg & Sartain, 2015). Teachers and administrators agreed that successful implementation of *Evaluate Davis* is largely dependent on principals' beliefs and attitudes about teacher evaluation in general, and specifically, their skillful use of the components of the evaluation tool. Principals would benefit from ongoing training to help them become more proficient and consistent evaluators (Goe, 2013). District leaders should work closely with school administrators to assure that principals have the knowledge and skill to be effective evaluators. For instance, teachers reported in the interviews that although they find the summative role of *Evaluate Davis* to be threatening, they also recognized the necessity of having an evaluation system that informs personnel decisions. Teachers and administrators agreed that having one evaluation system that fills both summative and formative roles adequately is difficult. Principals must become ultra-skilled in how they traverse the often-rocky terrain of dual-

purpose evaluation systems.

Environment. Principals should receive ongoing training in how to create an environment of accountability without the fear and distrust that teachers reported accompany summative evaluation. They should learn how they could assure that evaluations take place in an environment of support, collaboration, reflection, and professional development. Principals must become experts at developing and maintaining positive relationships with teachers and help them see evaluation as fair and useful (Danielson & McGreal, 2000; Havens & Roy, 2007; Marshall, 2005; Marzano & Toth, 2013; Popham, 2013; Toch, 2008; Towe, 2012; Weems & Rogers, 2010). Teachers in the interviews were clear that they feel they are driven to improve, that they reflect often on their teaching and look for ways to improve. What they hope for from administrators is encouragement and support in a friendly, growth-focused environment. They want to feel assured that their principals know what is going on in their classrooms and are ready and willing to take time with them to collaborate and develop improvement plans based on needs.

Feedback. Providing quality feedback is another skill principals need training on. Principals must become experts in gathering, analyzing, and interpreting data, and then presenting it to teachers in ways that motivate them to improve. Training should be required for principals to learn to give written and face-to-face feedback that focuses on strategies and solutions while acknowledging quality and growth. They must become competent in providing directed instructional assistance based on data they have gathered and information teachers provide. Principals need the knowledge and skill to develop a

school culture of industry and exploration that promotes frankness and confidence within the evaluation process (Darling-Hammond et al., 2012).

Time

Providing training and assurances to teachers that build confidence in principals' observations and evaluations is critical to the success of evaluation, but also requires more time than simply filling in checklists (Halverson et al., 2004). District leaders must work closely with principals to find the time to train teachers and to be trained themselves; to be in classrooms often, more than just the three or four times they are required to; to complete observation protocols effectively; and to spend time providing quality feedback and having frequent instruction-based conversations with teachers (Reinhorn, Johnson, & Simon, 2017). The many roles school administrators play should be reassessed and prioritized. Perhaps routine tasks such as hall and lunch supervision could be performed by teachers or other adults in the building at little additional cost. Possibly the district could reprioritize to free up funds for administrative interns who could take on the roles of student discipline, 504 coordinator, or LEA. Principals may want to re-evaluate roles assigned to the school leadership team to assure time and resources are used for the best benefit of teachers and students and time is freed up for classroom observations and follow-up conversations.

Focus on What Works

Current teacher evaluation literature recommends multiple data sources and lines of evidence (Darling-Hammond et al., 2012). *Evaluate Davis*, like many contemporary

evaluation systems, requires evaluators to gather data from an increased number of principal's observations each year, yearly parent and student surveys, and measures of student achievement. Principals in the interviews, however, reported that they often find themselves hurrying through observations, looking at surveys only to find support for employment decisions, and going through the motions of collecting evidence of student achievement, just to get evaluations done by district imposed deadlines. Teachers in the interview also reported that principals tried to save time by spending little time in classrooms and giving minimal feedback or comments on the observation protocol. They also placed little trust in parent and student surveys and were inclined to discount them for a number of reasons, indicating that they would not likely make changes based on them. Like the teachers I interviewed, the principals acknowledged only slight confidence in parent and student surveys.

Both groups recommended simplifying the evaluation process and focusing on the elements that have the most impact - evaluators being in classrooms often and following up with face-to-face conversations focused on student learning (Danielson, 2015). Both interview groups reported the belief that, of all the components of the evaluation process, conversations between administrators and teachers had the most powerful influence on teachers' growth and improvement and the greatest likelihood to improve student learning. Frequent conversations focused on providing teachers with the feedback they require and want, they agreed, create an environment of mutual trust and fairness in which teachers are most likely to take responsibility for, and engage in, meaningful professional development and growth plans which will deepen their understanding of

teaching and learning. Streamlining the process to focus on the components that make a difference will potentially give principals more time to focus on conversations and feedback and allow teachers more time to focus on self-reflection and assessment, setting meaningful goals, creating a professional development plan, and following through.

District leaders should restructure the training school administrators receive regarding implementation of *Evaluate Davis* to focus at least as much on how they administer the evaluation process as they do on what they are expected to do. Principals in the interview noted that they feel unprepared to engage teachers in instruction-focused dialogue following observations, saying they don't have time or the expertise. They expressed a need to learn how to facilitate post observation conversations that encourage teachers to own the improvement process through self-reflection and assessment, goal setting, and creating a self-improvement plan. They must learn how to ask the right questions that will guide teachers in identifying areas of strength and weakness, and they need to be more than familiar with the resources teachers will need to create an improvement plan.

Build Relationships

The administrator and teacher groups each talked about the need for administrators to develop relationships of trust with teachers. Each group recommended that district leaders ought to train principals in the skill of relationship building. The interview groups concurred that positive relationships between teachers and administrators are fostered by principals who engage teachers in meaningful dialogue about instructional practice, work closely with individual teachers to identify areas of

strength and weakness, and then collaboratively develop a plan for improvement. Teachers said that the more principals came into their rooms to actually see what is going on there, and then take time to have meaningful conversations, the more trusting their relationships are. They also said that when the relationship was positive, they were more likely to perceive the evaluation process as fair and valuable to them; they were more motivated to engage in collaboration with their peers regarding instructional practice, to self-reflect, and to participate in professional development.

Notwithstanding recommendations in the literature (Marzano & Toth, 2013; Popham, 2013) that quality relationships between teachers and administrators are foundational to successful evaluations, principals reported that they received little or no training from the school district on how to develop or maintain positive relationships. On the contrary, principals reported that most of the training they received was focused on the mechanics of the evaluation. If *Evaluate Davis* is to achieve its stated purpose of teacher growth, then district leaders should work more closely with school administrators to train and mentor them to cultivate relationships of trust with teachers (Frontier & Mielke, 2016).

Interrater Reliability

Teachers reported being skeptical of administrators' ratings, at least in part, because the ratings they received were dependent on which administrator observed and evaluated them. Principals reported that they made efforts to assure that each member of their administrator team was on the same page with respect to the way they rated teachers; they tried to ensure that teachers would receive the same ratings regardless of

which administrator was evaluating them.

One suggestion that came out of the interviews was to have the administrator team observe teachers together and compare and discuss ratings afterward. With each teacher observation, the ratings, feedback, and conversations with teachers should become more consistent between administrators. One school reported that all three administrators observe each teacher together on the first observation of the year, and meet with each teacher after the observation. They complete the other observations during the year individually, but still meet as a team to discuss any concerns. Based on district generated observation item reports, this school's administrative team has a high interrater reliability score.

Interrater reliability may also be improved through the establishment of common language and rubrics. Although the training materials for *Evaluate Davis* include rubrics for each item on the observation checklist, when I asked principals if they used them, some weren't aware they existed. Teachers asked if rubrics were available to them, as well. Ongoing professional development with teachers and administrators regarding the training rubrics would assist school leaders in promoting the instructional strategies they look for when observing. Additional ongoing training with the rubrics would also provide a springboard for quality instruction-focused conversations that include detailed, descriptive feedback (Weisberg et al., 2009).

Evaluate Davis includes an online professional development site called edPLUS which provides teachers and administrators with many learning options for teachers. For each of the five areas of the evaluation, PLAN, TEACH, CHECK, ENVIRONMENT,

and PROFESSIONALISM, teachers can simply click on any of the four or five indicators within the five areas and find resources such as web pages, videos, books, and articles that will give them specific ideas and solutions, based on well researched comprehensive teaching frameworks. This resource has been operational for four years, and yet none of the teachers and only one of the principals I interviewed knew it existed. The school district should do more to make principals and teachers aware of the resources available to them; principals should also recognize the resources they have at their fingertips that could assist in improving interrater reliability.

Limitations and Further Research

As with all research, this exploratory study has limitations. The sample population was small, limited to six teachers and six administrators, from elementary, junior high, and high schools, within one suburban school district in the state of Utah. Consequently, results are limited to the case that was studied. Opinions and perceptions of teachers and school administrators were gathered primarily through focus group interviews. Only one male teacher, and only one female administrator participated in the interviews, which may limit the range of views expressed. Participants' willingness to join in the study may be an indication of strong feelings they hold about the evaluation process that may have produced data that does not reflect the views of the general population at large. The perceptions and views shared in the group interviews may have been skewed one way or the other by a tendency toward group consensus or the charisma of one or two group members. The data gathered through the document review and the focus groups may be

slanted by my own biases about the evaluation process. Although I made every effort to look for data supporting a variety of viewpoints regarding the evaluation process, I can't ignore the reality that I have my own lens that I look at data through. Furthermore, the trustworthiness of this study is limited by the quality of the questions used in the focus group interviews and by the interviewer's skill in conducting reliable interviews.

This was an exploratory study and has significant limitations, as noted above. Similar studies should be conducted in comparable school districts to gain a better understanding of administrators' and teachers' perceptions of the effectiveness of teacher evaluation respecting teacher improvement and student growth. School and district cases should be identified and studied in which teaching is improving and student achievement is increasing. Additionally, further research, both qualitative and quantitative, should be conducted in a variety of school districts with different demographics, to compare results. Teachers' and administrators' perceptions of the teacher evaluation process may be very different in small rural school districts, or in large inner-city districts. They may also be different in charter schools or private schools. This was a qualitative study. More could possibly be learned using quantitative data from teacher and administrator surveys, student achievement data, and school features and environment.

More research should be conducted that considers the differences in the quality of the evaluation process between schools and districts, that looks at the different characteristics of the evaluation tools, the implementation processes, and the differences in the school-level leaders who administer the evaluation. For example, how skilled is the principal at accurately assessing teacher effectiveness, providing quality feedback, and

conducting post-observation conferences? How able is he or she at stimulating high-level instructional dialogue that evokes teacher reflection and self-improvement? How skillfully does she or he train faculty and staff regarding evaluation procedures? How well is professional development linked to specific needs identified through the evaluation process? Other studies should be conducted to gain greater understanding of teachers' and administrators' perceptions of teacher evaluation and their impact on the effectiveness of the evaluation process. Additional studies could add to the development of implementation strategies that contribute to the assurance of a successful evaluation experience in schools and districts.

Conclusion

One clear outcome of this study is an understanding that principals' perceptions, attitudes, and beliefs about the teacher evaluation process impact how it is implemented, how teachers receive it, and whether teacher growth and improvement will result. Another outcome is the realization that teachers' perceptions, attitudes, and beliefs about the evaluation process also have a significant influence on how they engage in the process and if they will grow and improve in their practice as a result. Policy makers and stakeholders inside and outside the realm of education profess a belief in the potential of teacher evaluation as a key, if not the key, to an improved teaching force and increases in student achievement. My experience and what I learned from interviewing teachers and administrators for this study is that there is still some reluctance on the part of educators to accept the evaluation process as formative, as a means of improving schools and

teachers. I suggest that those in schools and districts who are implementing new evaluation systems, like *Evaluate Davis*, need to consider the readiness level of schools, of principals and teachers, first. The culture of schools, the perceptions of school leaders, and the beliefs of teachers about their ability to change and improve need to be addressed if the changes in evaluation practice are to be sustained over time with any degree of success.

District and school level administration play a crucial role in advancing the rationale for changing evaluation procedures and connecting them to teacher and school improvements that affect student learning in positive ways. Once school leaders' and teachers' beliefs, attitudes, and values are in place, and adequate supports are engaged, effective practice will be the natural outcome. Teachers will value the feedback they receive and will have the desire to make necessary changes. Teachers and administrators will come to seek out the professional development needed to address areas of improvement identified through the evaluation process.

Although administrators in Davis School District have received extensive training in the tenets of *Evaluate Davis* and are required to pass a test of their competency as evaluators, additional, continuous training is still needed to assure they possess the pedagogical, collaboration, and systems change skills that will help assure success. Finding and engaging in continual professional development that focuses as much on how they implement the evaluation procedures as on what procedures to implement will make school leaders more competent formative evaluators and should be standard operating procedure.

Changes to teacher evaluation in Davis School District, like similar changes in districts throughout the state, have been motivated by attempts to meet the guidelines of federal initiatives aimed at improving learning for all students. That purpose should be foremost in every educational leader's mind and drive their motivation for improving their evaluation practices.

REFERENCES

- Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher*, 36(7), 369-387.
- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13(4), 544-559.
- Becker, H. S. (1998). *Tricks of the trade: How to think of your research while you're doing it*. Chicago, IL: University of Chicago Press.
- Bloor, M., & Wood, F. (2006) *Keywords in qualitative methods: A vocabulary of research concepts*. Thousand Oaks, CA: Sage.
- Bobbitt, F. (1918) *The curriculum*. New York, NY: Houghton Mifflin Company
- Bogden, R. C., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods* (3rd ed.). Boston, MA: Allyn & Bacon.
- Bogdan, R. C., & Biklen, S. K. (2003). *Qualitative research for education* (4th ed.). Boston, MA: Allyn & Bacon.
- Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theories and methods*. Boston, MA: Allyn & Bacon.
- Borman, G. D., & Kimball, S. M. (2005). Teacher quality and educational equality: Do teachers with higher standards-based evaluation ratings close achievement gaps? *Elementary School Journal*, 106(1), 3.
- Bruce, C. (2008). Countering the effects of poverty on students. *Diverse Learning Communities Today*, National Association of Elementary School Principals, 1(1), 1-2.
- Buddin, R., & Zamarro, G. (2009). *Teacher qualifications and student achievement in urban elementary schools*. Santa Monica, CA: RAND.
- Cochran-Smith, M. & Power, C. (2010). New directions for teacher preparation. *Educational Leadership*, 67(8), 6-13.
- Covey, S. R. (2014). *The 7 habits of highly effective families*. New York, NY: St. Martin's Press.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.

- Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson Education.
- Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
- Cummings, C. (2002). *Teaching makes a difference* (3rd ed.). Taiwan: Teaching, Inc.
- Danielson, C. (2001). New trends in teacher evaluation. *Educational Leadership*, 58(5), 12-15.
- Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.
- Danielson, C. (2014). Connecting common core to teacher evaluation. *School Administrator*, 71(3), 30-33.
- Danielson, C. (2015). *Talk about teaching! Leading professional conversations*. Thousand Oaks, CA: Corwin.
- Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: ASCD.
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, 8(1), 1.
- Darling-Hammond, L. (2012). The right start: Creating a strong foundation for the teaching career. *Phi Delta Kappan*, 94(3), 8-13.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan*, 93(6), 8-15.
- Darling-Hammond, L., Wise, A. E., & Klein, S. P. (1999). *A license to teach: Raising standards for teaching*. San Francisco, CA: Jossey-Bass.
- Davis School District. (2014). *Evaluate Davis: Licensed educator manual*. Farmington, UT: Author.
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York, NY: Macmillan.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educational process*. Lexington, MA: Heath.
- Donaldson, M. (2009). *So long Lake Wobegon? Using teacher evaluation to raise teacher quality*. Washington, DC: The Center for American Progress.

- Donaldson, M. (2010). No more valentines. *Educational Leadership*, 67(8), 54-58.
- Donaldson, M. (2016). Teacher evaluation reform: Focus, feedback, and fear. *Educational Leadership*, 73(8), 72-76.
- Donaldson, M., & Donaldson, G., Jr. (2012). Strengthening teacher evaluation: What district leaders can do. *Educational Leadership*, 69(8), 78-82.
- Donaldson, M., & Papay, J. (2015). An idea whose time had come: Negotiating teacher evaluation reform in New Haven, CT. *American Journal of Education*, 122(1), 39-70.
- Duffett, A., Farkas, S., Rothertham, A. J., & Silva, E. (2008). Waiting to be won over: Teachers speak on the profession, unions, and reform. *Education Sector*. Education Sector Reports.
- DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing students achievement*. Bloomington, IN: Solution Tree.
- Dufour, R. & Marzano, R. J. (2011). *Leaders of Learning: How District, School, And Classroom Leaders Improve Student Learning*, by Solution Tree Press
- DuFour, R., & Mattos, M. (2013). How do principals really improve schools? *Educational Leadership*, 70(7), 34-40.
- Every Student Succeeds Act of 2015, Pub. L. No. 114-95 § 114 Stat. 1177 (2015-2016).
- Fisher, D., Frey, N., & Pumpian, I. (2012). *How to create a culture of achievement in your school and classroom*. Alexandria, VA: ASCD.
- Flick, U. & Gibbs, G. (2007). Analyzing qualitative data. *Designing Qualitative Research*. Los Angeles, CA: SAGE
- Frontier, T., & Mielke, P. (2016). *Making teachers better not bitter: Balancing evaluation, supervision, and professional growth*. Alexandria, VA: ASCD.
- Gardner, D. P. (1983). *A nation at risk*. Washington, DC: The National Commission on Excellence in Education, U.S. Department of Education.
- Giossos, Y., Koutsouba, M., Lionarakas, A., & Skavantzios, K. (2009). Reconsidering Moore's transactional distance theory. *European Journal of Open, Distance and E-Learning*, 12(2). Retrieved from <http://www.eurodl.org/index.php?article=374>
- Given, L. M. (Ed.) (2008). *The Sage encyclopedia of qualitative research methods*. (Vol.2, pp. 697-698) Thousand Oaks, CA: Sage.

- Gleave, D. (1997). Bifocals for teacher development and appraisal. *Journal of Curriculum & Supervision*, 12(3), 269-281.
- Glesne, C. (2011). *Becoming qualitative researchers: An introduction*. Boston, MA: Pearson Education.
- Goddard, R. D., Hoy, W. K., & Woolfolk, H. A. (2000). Collective teacher efficacy: Its meaning, measure, and effect on student achievement. *American Education Research Journal*, 37, 479-507.
- Goe, L. (2013). Can teacher evaluation improve teaching? *Principal Leadership*, 13(7), 24-29.
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Goens, G. A., & Lange, R. W. (1983). Supervision and evaluation: Distinctions with differences. *The Developer*, (January), 1-4.
- Goldhaber, D. (2002). The mystery of good teaching. *Education Next*, 2(1), 50-55.
- Goldhaber, D., Anthony, E., & Urban Institute, (2005). *Can teacher quality be effectively assessed? National Board certification as a signal of effective teaching*. Washington, DC: Urban Institute.
- Goldrick, L. (2002). *Improving teacher evaluation to improve teaching quality*. Retrieved from: <http://files.eric.ed.gov/fulltext/ED480159.pdf>
- Goodwin, B., & Hein, H. (2016). Evaluating and improving: not the same thing. *Educational Leadership*, 73(8), 83-84.
- Gratz, D. B. (2011). Performance pay: Path to improvement. *Kappa Delta Pi Record*, 47(4), 156-161.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Quarterly*, 82(4), 581-629.
- Gubrium, J. F. & Holstein, J. A. (2002). *Handbook of interview research: Context and method*. Thousand Oaks, CA: Sage.
- Haefele, D. L. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education*, 7(1), 21-31.

- Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing teacher evaluation systems: How principals make sense of complex artifacts to shape local instructional practice. In W. Hoy & C. Miskel (Eds.), *Research and theory in educational administration* (pp. 153-188). Greenwich, CT: Information Age.
- Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to student achievement*. New York, NY: Routledge.
- Havens, M., & Roy, J. (2007). Teacher evaluation from a choice theory perspective. *Journal of Adventist Education*, (Feb./Mar.), 42-47.
- Haycock, K. (1998). Good teaching matters... a lot. *Thinking K-16*, 3(2), 1-14.
- Hazi, H. M., & Rucinski, D. (2009). Teacher evaluation as a policy target for improved student learning: A Fifty-state review of statute and regulatory action since NCLB. *Education Policy Analysis Archives*, 17(5), 1-22.
- Hollway, W., & Jefferson, T. (2000). *Doing qualitative research differently: Free association, narrative and the interview method*. Thousand Oaks, CA: Sage.
- Holstein, J. A., & Gubrium, J. F. (2003). *Inside interviewing: New lenses, new concerns*. Thousand Oaks, CA: Sage.
- Hunter, M. (1984). Knowing, teaching, and supervising. In P. Hosford (Ed.), *Using what we know about teaching* (pp. 169-192). Alexandria, VA: ASCD.
- Hunter, M. (1986). *The Hunter model of clinical supervision: A practical guide for instructional supervision: A tool for administrators and supervisors*. Sacramento, CA: Association of California School Administrators.
- Jackson, A. W., & Davis, G. A. (2000). *Turning points 2000: Educating adolescents in the 21st century*. Williston, VT: Teachers College Press.
- Kane, T. J., Taylor, E. S., Tyler, J. H., & Wooten, A. L. (2011). Evaluating teacher effectiveness. *Education Next*, 11(3), 54-60.
- Khan, S., & VanWynsberghe, R. (2008, January). Cultivating the under-mined: Cross-case analysis as knowledge mobilization. *Forum Qualitative Sozialforschung/ Forum: Qualitative Social Research*, 9(1), Art. 34, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0801348>.
- Kimball, S. M., & Milanowski, A. (2009). Examining teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, 45(1), 34-70.

- Kliebard, H. M. (2004). *The struggle for the American curriculum: 1853-1958* (3rd ed.). New York, NY: Routledge/Falmer.
- Kraft, M. A., & Gilmour, A. F. (2017). *Revisiting the widget effect: Teacher evaluation reforms and the distribution of teacher effectiveness*. Providence, RI: Brown University.
- LaFee, S. (2014). Students evaluating teachers. *School Administrator*, 71(3), 16-25.
- Lakshmi, B. S. (2014). Reflective practice through journal writing and peer observation: A case study. *Turkish Online Journal of Distance Education*, 15(4), 189-204.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Maeroff, G. I. (1996). *Breaking ranks: Changing an American institution*. Reston, VA: NASSP.
- Marchiando, K. D. (2013). *The influence of administrator perceptions on the teacher appraisal process* (Unpublished doctoral dissertation). University of Houston, Houston, TX.
- Marder, M. (2012). Measuring teacher quality with value-added modeling. *Kappa Delta Pi Record*, 48(4), 156-161.
- Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan*, 86(10), 727-735.
- Marshall, K. (2009). *Rethinking teacher supervision and evaluation*. San Francisco, CA: Jossey-Bass.
- Marshall, C., & Rossman, G. B. (2006). *Designing qualitative research*. Thousand Oaks, CA: Sage.
- Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: ASCD.
- Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14-19.
- Marzano, R. J., Frontier, T., & Livingston (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: ASCD.
- Marzano, R. J. & Toth, M. D. (2013). *Teacher evaluation that makes a difference: A new model for teacher growth and student achievement*. Alexandria, VA: ASCD

- Mathers, C. & Olivia, M. (2008). Improving instruction through effective teacher evaluation: Options for states and districts. *TQ Research and Policy Brief*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- Maxwell, J. C. (2007). *The 21 irrefutable laws of leadership: Follow them and people will follow you*. Nashville, TN: Nelson.
- Melrose, S. (2009). Naturalistic generalization. In A. J. Mills, G. Eurepos, & E. Wiebe (Eds.) *Encyclopedia of case study research* (pp. 599-600). Thousand Oaks, CA: Sage.
- Mendro, R. L. (1998). Student achievement and school and teacher accountability. *Journal of Personnel Evaluation in Education*, 12(3), 257-267.
- Merriam, S. (1988). *Case study research in education: a qualitative approach*. San Francisco, CA: Jossey-Bass.
- Merriam, S. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.
- Mielke, P. G. (2012). *Investigating a systematic process to develop teacher expertise: A comparative case study*. Milwaukee, WI: Cardinal Stritch University.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.
- Moore, M. G. (1991). Distance education theory. *The American Journal of Distance Education*, 1(5), 1-8.
- Moore, M. G. (1993). *Theory of transactional distance*. In D. Keegan (Ed.) *Theoretical principles of distance education*. New York, NY: Routledge.
- Morgan, D. L. (1997). *Focus groups as qualitative research* (Vol. 16). Thousand Oaks, CA: Sage.
- National Association of Secondary School Principals (2006). *Breaking ranks: A field guide for leading change*. Reston, VA: Author.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Author.
- National Commission on Teaching and America's Future. (1996). *What matters most: Teaching for America's future*. Woodbridge, VA: Author.

- Norman, A. D., (2010). Assessing accomplished teaching: Good strides, great challenges. *Theory into Practice, 49*(3), 203-212.
- Odden, A., & Kelley, C. (2002). *Paying teachers for what they know and do: Newer and smarter compensation strategies to improve schools* (2nd ed.). Thousand Oaks, CA: Corwin.
- Olatoye, R. A., & Aanu, E. M. (2011). Senior secondary school science teachers' perception of using student to evaluate teaching effectiveness. *Journal of Emerging Trends in Educational Research & Policy Studies, 2*(3), 164-170.
- Pallas, A. M. (2010). Measuring what matters most. *Phi Delta Kappan, 92*(4), 68-71.
- Palys, T. (2008). Purposive sampling. In L. M. Given (Ed.) *The Sage encyclopedia of qualitative research methods*. (Vol. 2, pp. 697-698). Los Angeles, CA: Sage.
- Patton, M. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage.
- Peterson, K. (2004). Research on school teacher evaluation. *NASSP Bulletin, 88*, 60-79.
- Popham, W. J. (1988). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education, 1*(3), 269-273.
- Popham, W. J. (2013). On serving two masters: Formative and summative teacher evaluation. *Principal Leadership, 13*(9), 18-22.
- Porter-Magee, K. (2004). Teacher quality, controversy, and NCLB. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78*(1), 26-29.
- Prothero N. (2008). Teacher efficacy: What is it and does it matter? *Principal, 87*(5), 42-45.
- Readers Digest. (1947, September). Quote page, (Volume 51, p. 64), New York, NY: Author.
- Reinhorn, S., Johnson, S. M., & Simon, N. (2017). Investing in development: Six high-performing, high-poverty schools implement the Massachusetts teacher evaluation policy. *Educational Evaluation and Policy Analysis, 39*(3), 383-406.
- Rennert-Ariev, P. (2005). A theoretical model for the authentic assessment of teaching. *Practical Assessment Research & Evaluation, 10*(2), 151-163.
- Rivers, J. C., & Sanders, W. L. (2002). Teacher quality and equity in educational opportunity: Findings and policy implications. In L. T. Izumi & W. M. Evers (Eds.). *Teacher Quality* (pp. 13-23). Palo Alto, CA: Hoover Institution Press.

- Rogers, E. M. (1983). *Diffusion of innovations*, 1st ed. New York: Free Press
- Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement* (Research Progress Report). Knoxville, TN: University of Tennessee Value Added Research and Assessment Center.
- Sawchuck, S. (2016). Law could spur changes in teacher requirements. *Education Week*, 35(15), 14-15.
- Schmoker, M. (2012). The next education fad: Complex teacher evaluations that don't work. *Education Week*, 32(2), 20, 24.
- Schon, D. A. (1988). Coaching reflective teaching. In P. P. Grimmett & G. L. Erickson (Eds.), *Reflection in teacher education* (pp. 39-46). New York, NY: Teachers College Press.
- Shavelson, R. J., & Towne, L. (Eds.). (2002). *Scientific research in education*. Washington, DC: National Research Council, National Academy Press.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information*, 22(2), 63-75.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Stake, R. E. (2005). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 443-466). Thousand Oaks, CA: Sage.
- Stake, R. E. (2008). Qualitative case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of qualitative inquiry* (3rd ed. pp. 119-150) Thousand Oaks, CA: SAGE
- Stearns, P. (1993). *Meaning over memory: Recasting the teaching of history and culture*. Chapel Hill, NC: UNC Press.
- Steinberg, M. P., & Sartain, L. (2015). Does better observation make better teachers? *Education Next*, 15(1), 70-76.
- Sullivan, S., & Glanz, J. (2009). *Supervision that improves teaching and learning: Strategies and techniques* (3rd ed.). Thousand Oaks, CA.: Corwin.
- Teaching Commission. (2004). *Teaching at risk: A call to action*. New York, NY: Author
- Toch, T. (2008). Fixing teacher evaluation. *Educational Leadership*, 66(2), 32-37.
- Torff, B. (2011). Teacher beliefs shape learning for all students. *Phi Delta Kappan*, 93(3), 21-23.

- Towe, P. B. (2012). *An investigation of the role of a teacher evaluation system and its influence on teacher practice and professional growth in four urban high schools*. Seton Hall, MO: Seton Hall University.
- Tucker, P., & Stronge, J. (2005) *Linking teacher evaluation and student learning*. Alexandria, VA: ASCD.
- U.S. Center for Public Education. (2005). *High performing, high poverty schools: Research review*. Fairfax, VA: Author.
- U.S. Department of Education. (2012). *ESEA flexibility*. Retrieved from www.ed.gov/esea/flexibility
- Valente, T. W. & Rogers, E. M. (1995). The origins and development of the diffusion of innovations paradigm as an example of scientific growth. *Science Communication*, 16(242).
- Walsh, K. (2004). Through the looking glass: How NCLB's promise requires facing some hard truths about teacher quality. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 78(1), 22-25.
- Wang, M., Haertel, G., & Walberg, H. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63(3), 249-294.
- Weems, D. M., & Rogers, C B. H. (2010). Are U.S. teachers making the grade? A proposed framework for teacher evaluation and professional growth. *Management in Education*, 24(1), 19-24.
- Weisberg, D., Sexton, S., Mulher, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: New Teacher Project.
- Weiss, E. M., & Weiss, S. G. (1998). New directions in teacher evaluation. ERIC Digest, ED429052 1998-12-00.
- Wetzel, W. (1929). Scientific supervision and curriculum building. *The School Review*, 37(2), 179-192.
- Whitaker, T. (2004) *What great teachers do differently: 14 things that matter most*. Larchmont, NY: Eye On Education.
- Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Thousand Oaks, CA: Sage.
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.) Thousand Oaks, CA: Sage.

Zatynski, M. (2012). Revamping teacher evaluation. *Principal*, 91(5), 22-27.

Ziebarth-Bovill, J., Kritzer, J., & Bovill, R. (2012). The essential criteria for hiring first year teacher candidates. *Education*, 133(1), 125-138.

APPENDICES

Appendix A

Permission Letter from Davis School District to Use Evaluation Documents

Bryon,

You have permission to cite Davis School District evaluation documents in your dissertation.

Patti Brown, Director
Evaluation & Quality Staffing
Davis School District

Appendix B
Interview Protocol

Interview Protocol: Dissertation: *An Investigation of Educators' Perceptions of the Influence of a Teacher Evaluation System on Student Learning.*

Time of Interview:

Date:

Setting:

Participants:

Description: This study is being conducted as partial fulfillment of the requirements for an Ed.D degree at Utah State University, Department of Teacher Education and Leadership, and is a case study, in which the attitudes, beliefs, and perceptions of teachers and school administrators in the Davis School District in Davis County, Utah, are being examined in the interest of developing a greater conception of how their perceptions, attitudes, and beliefs influence the effectiveness of the current teacher evaluation tool, *Evaluate Davis*, as implemented in Davis School District.

Questions:

- Is *Evaluate Davis* helpful in improving the quality of teaching in your school?
- Has increasing the number of times teachers are observed each year been beneficial to teachers and students?
- How valid do you believe parent and student surveys are in terms of the feedback they provide teachers?
- Has the quality of feedback teachers receive been beneficial in terms of improving the quality of instruction students get?
- Considering provisional teachers, mid-career teachers, and teachers with 20 or more years of experience, which seem to be the most amenable to the feedback they receive from evaluators?
- How can *Evaluate Davis* be improved to achieve the overall goal of increased student learning?

Appendix C
Informed Consent Form

INFORMED CONSENT

An Investigation of Educators' Perceptions of the Influence of a Teacher Evaluation System on Student Learning.

Introduction

You are invited to participate in a research study conducted by Bryon Nielsen, a doctoral student in the

Department of Teacher Education and Leadership at Utah State University. The purpose of this research is to learn about educators' attitudes and perceptions of the yearly teacher evaluation process.

This form includes detailed information on the research to help you decide whether to participate in this study.

Please read it carefully and ask any questions you have before you agree to participate.

Procedures

Your participation will involve participation in an approximately 45-minute-long focus group interview in which you will be asked some questions about your experience with the evaluation process being investigated.

Responses will be recorded and data gathered to inform the research study. There are no right or wrong answers to the questions; they are asked to get your impression of the evaluation process. Your total participation will be no more than 45 minutes.

If you agree to participate, the researchers will also collect pertinent demographic data such as the number of years you have been an educator, what subject(s) you teach, your gender, and the age level you teach (high school, junior high, or elementary). We anticipate that 12 people will participate in this research study.

Risks

This is a minimal risk research study. That means that the risks of participating are no more likely or serious than those you encounter in everyday activities. Participation may involve some added risks or discomforts. These are minimal but there is a small chance confidentiality could be breached. Bryon Nielsen will take steps to reduce the risk. This study will be used specifically for the purpose of gathering information for a research project and will not be released to the school district.

Benefits

There is no direct benefit to you for participating in this research study. More broadly,

this study will help the researcher learn more about teachers' and administrators' attitudes and perceptions of the evaluation process and may result in improvements to the evaluation process and thus provide indirect benefits now and in the future.

Voluntary nature of participation and right to withdraw without consequence

Your participation in this research is completely voluntary. If you agree to participate now and change your mind later, you may withdraw at any time by simply contacting Bryon Nielsen at 801-231-6815 or bryonnielsen@gmail.com. If you choose to withdraw after we have already collected information about you, to the extent possible, your information will be removed from the study and destroyed. (Completely anonymous participation cannot be withdrawn as the researcher will be unable to determine whose data is whose.) There are no consequences for deciding to withdraw from the study.

Payment or Compensation

There is no compensation or payment for participating in this study.

Confidentiality

The researchers will make every effort to ensure that the information you provide as part of this study remains confidential. Your identity will not be revealed in any publications, presentations, or reports resulting from this research study. However, it may be possible for someone to recognize your particular story/situation/response (particularly applicable in focus group/ethnographic/oral history research projects). While we will ask all group members to keep the information they hear in this group confidential, we cannot guarantee that everyone will do so.

We will collect your information through focus group interviews which will be digitally recorded. Data will be securely stored in a restricted-access folder on Box.com, an encrypted, cloud-based storage system and locked in a restricted-access office. To protect your privacy, personal, identifiable information will be removed from documents and replaced with a study identifier. Identifying information will be stored separately from data and will be kept until data collection is complete. Audio recordings will be destroyed after transcriptions are complete.

This study will be completed in May, 2017 and all personally identifiable information will be destroyed three years from that date. This form will be kept for three years after the study is complete, and then it will be destroyed.

It is unlikely, but possible, that others (Utah State University, or state or federal officials) may require us to share the information you give us from the study to ensure that the research was conducted safely and appropriately.

We will only share your information if law or policy requires us to do so.

IRB Approval Statement

The Institutional Review Board (IRB) for the protection of human research participants at Utah State University has reviewed and approved this study. If you have questions about the research study itself, please contact the

Principal Investigator at (801) 231-6815 or bryonnielsen@gmail.com. If you have questions about your rights or would simply like to speak with someone other than the research team about questions or concerns, please contact the IRB Director at (435) 797-0567 or irb@usu.edu.

Informed Consent

By signing below, you agree to participate in this study. You indicate that you understand the risks and benefits of participation, and that you know what you will be asked to do. You also agree that you have asked any questions you might have, and are clear on how to stop your participation in the study if you choose to do so. Please be sure to retain a copy of this form for your records.

Participant's Signature	Participant's Name, Printed	Date
-------------------------	-----------------------------	------

Investigator Statement

“I certify that the research study has been explained to the individual, by me or my research staff, and that the individual understands the nature and purpose, the possible risks and benefits associated with taking part in this research study. Any questions that have been raised have been answered.”

Signature of Researcher(s)

Dr. Courtney Stewart
Principal Investigator
(435) 797-7145
Courtney.stewart@usu.edu

Bryon Nielsen
Student Researcher
801-231-6815
bryonnielsen@gmail.com

Appendix D

PLAN Rubric

STANDARD & INDICATOR	CLASSROOM OBSERVATION MEASURE	WHY IS IT IMPORTANT?	WHAT EVIDENCE MIGHT I SEE?	RED FLAGS
<p>PLAN</p> <p>P.1 Analyzes and uses ongoing assessment data to guide instructional planning</p>	<p>1. Not part of the observation measure</p>	<p>Effective teachers constantly monitor student learning using a variety of both formative and summative information. Analyzing and using data enables teachers to make data-based instructional decisions, and take students from where they are to where they need to be.</p>	<ul style="list-style-type: none"> Teachers reviewing data as a guide to plan instruction. Teachers collaborating with peers in data teams and other assessment results to guide their discussions and planning for instruction. Teacher refers to data when sharing with students why or how they are learning what they are learning. Teachers using or referring to data boards when discussing instructional strategies and lesson planning. A quick DESK check enables observer to identify the standard/ indicator being taught. 	<ul style="list-style-type: none"> Teachers does not know how well a student is doing in the class. Students either do not have enough background information to use the new information or the concept being taught is something they already know.
<p>P.2 Designs instruction to meet or exceed DESK standards</p>	<p>2. Lesson aligns with DESK standards</p>	<p>The DESK identifies what our students should know and be able to do. It is a prioritized and focused set of standards aligned with the Utah Core. Because our time with students is limited, lessons should focus on ensuring our students can meet these standards.</p>	<ul style="list-style-type: none"> Observer cannot relate lesson to any item in DESK. Inordinate amount of time spent doing "housekeeping" or discussing topics not related to curriculum. 	<ul style="list-style-type: none"> Worksheets, crossword puzzles and word searches. Creative projects that are not related to the DESK. Students are unsure of what they are supposed to be learning from the activity or assignment – completion is the goal.
<p>P.3 Constructs relevant meaningful learning experiences which meet individual learning needs</p>	<p>3. Activities help learners master content or skills</p>	<p>Many activities are fun or interesting (or not) but have minimal impact on student learning. Because we have limited time with students, the learning activities must effectively help learners master the content or skill.</p>	<ul style="list-style-type: none"> Students actively working individually or in groups. Students asking questions about skills or content rather than procedural concerns. Teacher asks students, "Why are we doing this?" or "How will this activity help us to master _____?" or "Why is this activity important?" 	<ul style="list-style-type: none"> Teachers use skills in isolation and do not have opportunities to relate it to other situations, ideas, or use it in any other way. Material is 'covered' and then the teacher moves on. There is no 'practice' or application of the material or skills. Students cannot answer the question, "How will you use this?"
<p>4. Teacher provides opportunities for learners to apply information to new situations</p>	<p>4. Teacher provides opportunities for learners to apply information to new situations</p>	<p>Learning knowledge and mastering skills is important, however, more important is that learners are able to take the information or skill they have learned and apply it in new or different situations. When they are able to generalize information to new situations and make connections between what they are learning in class and what they come in contact with in other content area learning or learning outside the classroom, learning is deepened and retained.</p>	<ul style="list-style-type: none"> Teacher and students use statements like, "Now, let's try it in this situation," or, "We talked about how we might do this for this assignment, but how would we use this if we were at _____ doing _____" or teacher or students bring up current events and relates to learning concepts. 	<ul style="list-style-type: none"> All students are engaged in doing the same activity in the same way. Some students are struggling and unable to complete assignment. Some students finish quickly and are doing other work (peer tutoring, reading a book, etc.) Students can't relate content or skill to
<p>5. Activities are differentiated to meet individual needs of learner</p>	<p>5. Activities are differentiated to meet individual needs of learner</p>	<p>Learning experiences are meaningful and relevant when they meet individual learning needs. Knowing where students are in terms of their skill and ability levels, and their interests enables an educator to construct learning experiences that engage them, hold their interest, and enable them to be successful. All students, including advanced</p>	<ul style="list-style-type: none"> Most students complete the assignment in approximately the same amount of time. Some students have modified assignments. Students share comments on how the concept relates to their interests or occurrences in their lives. 	<ul style="list-style-type: none"> Students cannot answer the question, "How will you use this?"

Appendix E

Administrator and Teacher Interview Schema
Document Review Schema

Administrator Interview Schema

Pseudonym	Research Question 1	Research Question 2	Research Question 3
Marshall			
Seth			
Dan			
Brad			
Trent			
Jill			

Teacher Interview Schema

Pseudonym	Research Question 1	Research Question 2	Research Question 3
Karen			
Nadine			
Loraine			
Christy			
Tara			
Tom			

Document Review Schema

Document	Research Question 1	Research Question 2	Research Question 3
Rubrics			
Power Point Slides			
Leader Notes			
Emails			
New Teacher Training			
All Faculty Training			

Appendix F

Links to Effective Utah Teaching Standards and *Evaluate Davis* Manuals

Utah Effective Teaching Standards

<http://www.schools.utah.gov/CURR/educatoreffectiveness/Standards/Teaching/TeachingStandards.aspx>

Evaluate Davis Homepage

<https://www.davis.k12.ut.us/domain/9322>

CURRICULUM VITAE

BRYON NIELSEN

EDUCATION

- Doctor of Education, Curriculum and Instruction, Utah State University, 2018
- Master of Science, Special Education, University of Utah, 1986
- Bachelor of Science, Psychology, University of Utah, 1981

EXPERIENCE

- Math and science teacher, Preston High School, Preston, Idaho 1981-1982
- Math and science teacher, Bountiful Junior High School, Bountiful, Utah 1982-1988
- Math and science teacher, Davis Learning Center (self-contained school for students with severe behavior disorders), Layton, Utah 1988-1991
- Assistant Principal, Millcreek Junior High School, Bountiful, Utah 1991-1995
- Principal, Millcreek Junior High School, Bountiful, Utah 1995-2003
- Principal, South Davis Junior High School, Bountiful, Utah 2003-2011
- Principal, Fairfield Junior High School, Kaysville, Utah 2011-present

PUBLICATIONS

- Helping All Students Achieve Their Potential, *Impact Journal*, Utah Association of Secondary School Principals, (2015), Volume 16, Issue 1.

RECOGNITION

- Utah Principals Academy Fellow 1998-1999
- Administrator of the Year, Utah Educational Library Media Association 1998-1999
- Secondary Principal of the Year, Davis School District, 2001

PROFESSIONAL ASSOCIATIONS

- National Association of Secondary School Principals
- Utah Association of Secondary School Principals
- Association for Supervision and Curriculum Development
- Utah Association for Supervision and Curriculum Development
- Association for Middle Level Education
- Utah Middle Level Association

PROFESSIONAL PRESENTATIONS:

Utah Middle Level Association Annual Conference, March 14, 2015