Principals’ Perceptions of Their Own Evaluation Process: How Feedback Using the Colorado Model Evaluation System Alters Principals’ Practice

Timothy Kevin Ridder

Follow this and additional works at: https://digscholarship.unco.edu/dissertations

Recommended Citation

This Text is brought to you for free and open access by the Student Research at Scholarship & Creative Works @ Digital UNC. It has been accepted for inclusion in Dissertations by an authorized administrator of Scholarship & Creative Works @ Digital UNC. For more information, please contact Jane.Monson@unco.edu.
PRINCIPALS’ PERCEPTIONS OF THEIR OWN EVALUATION PROCESS: HOW FEEDBACK USING THE COLORADO MODEL EVALUATION SYSTEM ALTERS PRINCIPALS’ PRACTICE

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Education

Timothy Kevin Ridder

College of Education and Behavioral Sciences
Department of Leadership, Policy and Development:
Higher Education and P-12 Education
Educational Leadership and Policy Studies Program

August 2018
This Dissertation by: Timothy Kevin Ridder


Has been approved as meeting the requirement for the Degree of Doctor of Education in the College of Education and Behavioral Sciences, Educational Leadership and Policy Studies Program

Accepted by the Doctoral Committee

__________________________________________________________
Spencer Weiler Ph.D, Research Advisor

__________________________________________________________
Amie Cieminski Ed.D, Committee Member

__________________________________________________________
Jacob Skousen Ed.D, Committee Member

__________________________________________________________
Jenni Harding Ed.D, Faculty Representative

Date of Dissertation Defense: March 28, 2018

Accepted by the Graduate School

__________________________________________________________
Linda L. Black, Ed.D.
Associate Provost and Dean
Graduate School and International Admissions
ABSTRACT


**Background:** The Colorado General Assembly created the Educator Effectiveness Act in 2011. This legislation established the State Council for Educator Effectiveness (The Council). The Council was tasked with developing the purpose behind and structure for a new evaluation system for the state of Colorado. The purpose of the Colorado Model Evaluation System (CMES) established by the Council was “to provide meaningful and credible feedback that improves performance” (State Council for Educator Effectiveness, 2011). With the purpose of and outline for the new evaluation system created by the Council, the Colorado Department of Education developed the Colorado Model Evaluation System (CMES). The CMES was initially piloted and fully adopted during the 2014-2015 school year. One population of educators being evaluated using the CMES was principals. **Purpose:** The purpose of this research study was to identify principals’ perceptions of the feedback they were receiving as a part of the CMES. The research questions for this study were: What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process? What are principals’ perceptions of the CMES evaluation feedback? According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice? Participants for the study were principals being evaluated using the CMES. The method
used to collect the research was an online survey questionnaire using Qualtrics. The research was analyzed using quantitative, qualitative, and mixed-methods strategies.

**Findings:** There were 152 participants in this study. The largest number of participants perceived the CMES as positively altering their practice. The second largest number of participants believed the feedback had no impact on their practice. The main positive perceptions of the CMES and feedback were that the system helps them focus their attention on their practice and that it is a system based on improvement. The perceived areas where CMES and feedback can improve were by making the system less cumbersome, and the potential improved connection between the evaluators, the system, and the principals being evaluated.

**Key Words:** principal evaluation feedback principal practice
ACKNOWLEDGEMENTS

Thank you to my research committee, most especially Dr. Spencer Weiler, who became my advisor at a time when I was ready to quit and saw me through to the end. To my mom and dad who gave me plenty pep-talks throughout this process. To my brothers and sisters who always checked in with me, especially my sister Sarah who helped me early on whenever she could. To my children with whom I need to make up many years of play time. To my wife who tolerated the times I was away and working, and the wonderful energy she gave me that I needed from beginning to end.

In my first semester, we read an article that told all the ways that negatively impact the doctoral process and ultimately cause people to fail. After six plus years I managed to do everything that was supposed to make me quit and I still made it. Thank you for being done.
# TABLE OF CONTENTS

## I. INTRODUCTION ................................................................. 1
- Background of the Study – Government Directives
- Background – Theories on Feedback
- Statement of the Problem
- Research Question
- Definition of Key Terms
- Conclusion

## II. REVIEW OF LITERATURE.................................................. 14
- Mediating Variables: Principal Impact
- Mediating Variables: Standards, Role Conceptions, Skills, Responsibilities, and Character Traits of Principals
- Principal Evaluations
- Role(s) of the Evaluators of Principals
- Research on Feedback Intervention Theory and Regulatory Focus
  - Theory
  - Gap in the Research

## III. METHODOLOGY .............................................................. 72
- Overview
- Research Question
- Research Design
- Researcher Bias
- Assumptions of This Study
- Conclusion

## IV. ANALYSIS ........................................................................... 114
- Survey Response and Demographics Analysis
- Answers to the Research Questions
- Conclusion to Chapter IV

## V. CONCLUSIONS AND RECOMMENDATIONS............................. 169
- Response Rate
- Discussion of the Three Research Questions
- Implications
- Limitations of the Study
- Recommendations for Future Research
- Recommendations for the Colorado Department of Education
- Conclusion
REFERENCES ........................................................................................................... 186

APPENDIX

A. FRAMEWORK FOR SYSTEM TO EVALUATE PRINCIPALS............................................................... 198

B. STATE SCORING FRAMEWORK MATRIX FOR PRINCIPALS............................................................ 200

C. TABLE C.1: COLORADO DEPARTMENT OF EDUCATION’S SURVEY RESULTS OF PILOT PRINCIPALS’ PERCEPTIONS........... 202

D. FINAL SURVEY FOR PRINCIPALS.......................................................................................... 204

E. CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH – UNIVERSITY OF NORTHERN COLORADO................................................................. 216

F. UNIVERSITY OF NORTHERN COLORADO INSTITUTIONAL REVIEW BOARD APPROVAL................................. 219

G. INITIAL EMAIL SENT TO ALL PRINCIPALS................................................................. 221

H. TABLES H.1, H.2, & H.3 – CROSSTABS COMPARING FREQUENCY OF FEEDBACK WITH THE THREE RESEARCH QUESTIONS........................................................................... 223

I. TABLE I.1: QUALITATIVE CATEGORIES FOR ALL THREE RESEARCH QUESTIONS.............................. 227
LIST OF TABLES

Table 2.1  Principals’ Standards, Role Conceptions, Skills, and Responsibilities Found by Researchers……………………………... 23
Table 4.1  Years participants have been principals……………………………………….. 121
Table 4.2  Years participants have been in current position…………………………….. 122
Table 4.3  Location of school district in Colorado……………………………………….. 122
Table 4.4  Type of school district: State of Colorado verses participants……… 123
Table 4.5  FARM rate: state of Colorado verses participants…………………………….. 123
Table 4.6  Size of School: State of Colorado Verses Participants………………… 124
Table 4.7  Type of school: State of Colorado Verses Participants………………… 125
Table 4.8  Final Effectiveness Rating – Sample Verses State of Colorado……… 125
Table 4.9  How Often Do You Receive Evaluative Feedback for the Colorado Model Evaluation System……………………………………… 127
Table 4.10 What is Your Overall Perception of the Colorado Model Evaluation System for Principals……………………………………… 129
Table 4.11 Frequency of Feedback Compared to the Answers to Research Question Number One……………………………………… 131
Table 4.12 What Has Been Your Perception of the Feedback You Have Received as a Part of the Colorado Model Evaluation System…….. 144
Table 4.13 Frequency of Feedback Compared to the Answers Given for Research Question Number 2……………………………………… 146
Table 4.14 How Has Feedback You Have Received Using the Colorado Model Evaluation System Changed Your Work as a Principal….. 157
Table 4.15 Frequency of Feedback Compared to the Results for Research Question Number Three……………………………………… 158
Table C.1  Percent of Positive Responses Given by Principals in Baseline and Feedback Surveys

Table H.1  Frequency of feedback using the CMES and principal perception of CMES – Crosstab

Table H.2  Frequency of feedback using the CMES and principal perception of feedback using CMES – Crosstab

Table H.3  Frequency of feedback using the CMES and principal perception of how CMES has changed principal practice – Crosstab

Table I.1  Themes for all three research questions
LIST OF FIGURES

Figure 2.1  Principals Effects Models Adopted from Pitner
(as cited in Hallinger & Heck, 1996, p. 16)………………….. 16

Figure 2.2  This graphic demonstrates the complicated impact of the principal
on student achievement (Ahtaridou et al., 2009, p. 191)………… 20

Figure 2.3  Visual Model of CMES process for all educators……………… 44

Figure 4.1  Data analysis process for this mixed methods study……………….. 98
CHAPTER I
INTRODUCTION

During the 2014 – 2015 school year a new evaluation system for educators was fully implemented in the state of Colorado (Executive Order B 2010-001, 2010; C.R.S. 22-9-105.5, 2013). There were five assumptions used when creating this system:

1. Data should inform decisions, but human judgment will always be an essential component of evaluation,
2. The implementation and assessment of the evaluation system must always embody continuous improvement,
3. The purpose of the system is to provide meaningful and credible feedback that improves performance,
4. The development and implementation of educator evaluation systems must continue to involve all stakeholders in a collaborative process, and
5. Educator evaluation systems must take place within a larger system that is aligned and supportive. (State Council for Educator Effectiveness, 2011, pp. 5-8)

Of particular interest to this study was the third assumption, which makes clear, that the purpose of the Colorado Model Evaluation System (CMES) for the educators and evaluators in the state of Colorado is to provide all educators with meaningful feedback. Meaningful and credible feedback should be provided to help educators improve their performance (State Council for Educator Effectiveness, 2011). The researcher of this
dissertation studied the perceptions of feedback received by Colorado principals that were evaluated using the CMES.

Including this introduction, there are seven sections to this chapter. In the section that follows, the researcher outlines the federal and state directives that led to the CMES. In the third section a theoretical perspective is given about feedback through the explanation of two theories. The fourth section is the Statement of the Problem. The researcher then briefly introduces the research question in the next section. In the sixth section, key terms are defined. The final section is the conclusion of this chapter. Overall, the researcher will provide a foundation of understanding for this research.

**Background of the Study – Government Directives**

There are two parts to the background of this study. The first part is the legislation that led to the CMES. The second part provides a theoretical perspective of feedback. For the initial background of this study, the story of the legislation that led to the CMES will be told. First, *Race to the Top*, the federal grant that initiated the most recent push for changes in educator evaluations, will be explained. Second, Colorado Governor Bill Ritter’s executive order, *Executive Order B 10 001* will be outlined. To conclude this section, the Colorado General Assembly’s legislation, *Educator Effectiveness Act* will be detailed.

**Race to the Top**

As the economy was struggling in the latter part of the 2000s, the Obama administration put together a comprehensive bill to stimulate the economy. “On February 17, 2009, President Obama signed into law the American Recovery and Reinvestment Act of 2009 (ARRA), historic legislation designed to stimulate the
economy, support job creation, and invest in critical sectors, including education” (US Department of Education, 2009, p. 1). One part of the grant was *Race to the Top* (RttT) which was a “competitive grant program designed to encourage and reward states that are creating the conditions for education innovation and reform” (US Department of Education, 2009, p. 2). One of the four foci for RttT was “recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most” (US Department of Education, 2009, p.2). The following points demonstrate that this focus was very important to the Obama administration to improve schools.

Before state officials could submit the RttT application, they had to meet two requirements (US Department of Education, 2009). The first was that the application had to be approved by the US Department of Education before it was accepted (US Department of Education). The second condition was that, “there must not be any legal, statutory, or regulatory barriers at the State level to linking data on student achievement or student growth to teachers and principals for the purpose of teacher and principal evaluation” (US Department of Education, p. 4 (b)). The second pre-application requirement emphasized the need to link evaluations to student achievement (US Department of Education).

After passing the initial requirements, state officials went through an application process in which their applications were scored using a rubric that was developed by the US Department of Education (US Department of Education, 2009). States were rewarded points for each section of their application (US Department of Education). The US Department of Education gave the criterion called “Great Teachers and Leaders” the
highest point total with 138 points out of a total of 485 points. Under the “Great Teachers and Leaders” criterion, there were five sub-criteria. Of the five sub-criteria, the one with the highest point total within Great Teachers and Leaders is “Improving teacher and principal effectiveness based on performance (58 points)” (US Department of Education, p. 3). The US Department of Education (2009) prioritized teacher and principal performance by assigning this sub-criterion the highest point total.¹

The expectations set for the sub-criterion “Improving teacher and principal effectiveness based on performance” included clear approaches to student growth, designing evaluations systems that use multiple rating categories that include student growth, “conduct annual evaluations of teachers and principals that include timely and constructive feedback” (US Department of Education, 2009, p. 3). The use of evaluations to inform decisions on compensation, promoting retaining principals and teachers, tenure decisions, and removal of ineffective teachers and principals were supposed to be a priority for states officials vying for RttT funds (US Department of Education).

These two factors suggest that part of the priorities of the federal government was to reform the teacher and principal evaluation in the Race to the Top competition by the US Department of Education (2009). First, to apply for RttT funding there could not be any barriers to linking student data to teacher and principal evaluations and, second, the number of points assigned specifically to the criterion associated with reforming

¹ The other sub criterion for “Great Teachers and Leaders” were: Providing high-quality pathways for aspiring teachers and principals (21 points), Ensuring equitable distribution of effective teachers and principals (25 points), Improving the effectiveness of teacher and principal preparedness programs (14 points), and Providing effective support to teachers and principals (20 points) (US Department of Education, 2009, p. 3).
evaluation surpassed any other criterion (US Department of Education, 2009). In 2009, the federal government established its national interest in education through its attempt to reward the reform of teacher and principal evaluation through RttT funds (US Department of Education, 2009). The State of Colorado was one of the states to apply for these funds (Executive Office of the State of Colorado, 2010).

**Executive Order B 2010-001 and the Educator Effectiveness Act**

As the state of Colorado made promises of reform in an application to the federal government for RttT funding, Governor Bill Ritter and the Colorado General Assembly established change in Colorado’s education system through *Executive Order B 2010-001* and *The Educator Effectiveness Act* (Executive Order B 2010-001, 2010; C.R.S. 22-9-105, 2013). Through *Executive Order B 2010-001*, Governor Ritter set the stage for the *Educator Effectiveness Act*.

The *Educator Effectiveness Act* created the State Council for Educator Effectiveness (the Council) and established the purpose of the Council’s work, which was to ensure all licensed personnel are:

1. Evaluated using multiple fair, transparent, timely, rigorous, and valid methods, at least fifty percent of which evaluation is determined by the academic growth of their students;
2. Afforded a meaningful opportunity to improve their effectiveness; and
3. Provided the means to share effective practices with other educators throughout the state. (C.R.S. 22-9-105, 2013, (c) I-III).

After establishing the purpose, the authors of the *Educator Effectiveness Act* then described all the duties and responsibilities of the Council (C.R.S. 22-9-105, 2013). The
The result of the Council’s efforts was to be an evaluation system that was initiated during the 2011-2012 school year, piloted starting the 2012-2013 school year, and fully implemented in the 2014-2015 school year (C.R.S. 22-9-105, 2013).

The outline of Executive Order B 2010-001 and the Educator Effectiveness Act provides a background to the foundation of the CMES. The recommendations of the Council and an explanation of the CMES will be provided in the Literature Review of this study. In their recommendations, the Council stated that the purpose of the CMES is to provide feedback to improve performance (State Council for Educator Effectiveness, 2011). A second piece of the background for this study is the theoretical perspective of feedback given by the authors of Feedback Intervention Theory (FIT) and Regulatory Focus Theory (RFT) (Brockner & Higgins, 2001; Kluger & DeNisi, 1996).

**Background – Theories on Feedback**

Two theories address effective feedback interventions and the mentality of those receiving feedback (Brockner & Higgins, 2001; Kluger & DeNisi, 1996). Kluger and DeNisi (1996) created the Feedback Intervention Theory (FIT) to craft a consistent process to think of feedback interventions. Brockner and Higgins (2001), the authors of Regulatory Focus Theory (RFT), addressed the potential mentality those receiving feedback might be in and how they might react to certain types of feedback. Both theories provide a framework for delivering and receiving effective feedback.

**Feedback Intervention Theory**

Kluger and DeNisi (1996) established a theory of effective feedback interventions known as Feedback Intervention Theory (FIT). FIT is based on the following five arguments:
1. Behavior is regulated by comparisons of feedback to goals or standards,
2. Goals or standards are organized hierarchically,
3. Attention is limited and therefore only feedback-standard gaps that receive attention actively participate in behavior regulation,
4. Attention is normally directed to a moderate level of hierarchy, and
5. FIs [Feedback Interventions] change the locus of attention and therefore affect behavior. (Kluger & DeNisi, 1996, p. 259)

With FIT, if feedback interventions are purposeful, and follow the logic behind these five arguments, Kluger and DeNisi (1996) concluded that effective feedback interventions would be given. Kluger and DeNisi first argued that people will adjust their effort towards completing goals or standards per whether the feedback intervention is identifying a gap in performance. If the feedback intervention signals that the performance falls short compared to the standard, “effort is typically increased;” if the feedback intervention signals that the performance exceeds the standards, “effort is typically reduced (or maintained)” (Kluger & DeNisi, 1996, p. 260). Their second argument then organized the standards or goals per whether the goals or standards are focused on regulating peoples’ perceptions of themselves (self) or their performance on a specific task. Those feedback loops that are, “at the top of the hierarchy contain goals of the self, whereas those at the bottom of the hierarchy contain physical action goals” (Kluger & DeNisi, 1996, p. 261).

After identifying that there is a continuum of types of feedback interventions, Kluger and DeNisi (1996) then discussed that people focus on that which they pay
attention. The attempt at behavior regulation is limited by the feedback [intervention] – standard gaps that people prioritize with their focus. The most effective feedback interventions, which is the moderate level, were identified in the fourth argument. Attention should not be directed “to the ultimate goals of the self or to the detailed components of an ongoing activity” (Kluger & DeNisi, 1996, p. 262). To focus on either end of the hierarchy too heavily, can negatively impact performance. Kluger and DeNisi’s final argument made a final connection between feedback interventions, attention, and behavior. They stated since feedback interventions are given significant attention, and their purpose is to alter behavior to align with standards or goals, just by prioritizing attention toward feedback interventions, they will impact behavior. Kluger and DeNisi (1996) concluded that for feedback interventions to be successful, one must be conscious of these five arguments (Feedback Intervention Theory). If the goal of the CMES is to provide formative feedback to improve performance, then it could be valuable for principals and their evaluators using the CMES be conscious of these five arguments (State Council for Educator Effectiveness, 2011).

**Regulatory Focus Theory**

Whereas Kluger and DeNisi (1996) focused on feedback interventions and where these interventions should focus on the continuum from self to task in FIT, the authors of Regulatory Focus Theory (RFT) concentrated on the emotional state of people when they interact with their environments (Brockner & Higgins, 2001). Brockner and Higgins (2001) explained people’s emotional focus with two distinct categories: Promotion focused and Prevention focused. They then explained three factors that impact whether a person is promotion or prevention focused: “(a) the needs people are seeking to satisfy,
(b) the nature of the goal or standard that people are trying to achieve or match, and (c) the psychological situations that matter to people” (Brockner & Higgins, 2001, p. 37). The two categories and three factors are described in the paragraphs below.

Promotion focused people “are motivated by growth and development needs in alignment with their ideal selves” (Brockner & Higgins, 2001, p. 35). On the contrary, people who are prevention focused “are responsive to security needs in which they try to match their actual selves with their ought selves (self-standards based on felt duties and responsibilities)” (Brockner & Higgins, 2001, p. 35). To find if people are prevention or promotion focused will help to identify how to help them grow.

When explaining the impact of authority on RFT in those who would work for them, Brockner and Higgins (2001) suggested that people in authority are impactful on the regulatory focus of their work environments. “Authorities may affect subordinates’ tendencies to be promotion or prevention focused” (Brockner & Higgins, 2001, p. 60). Supervisors and evaluators should be conscious of their feedback environment and its impact on those who work for them.

There are three factors that construct people’s regulatory focus (Brockner & Higgins, 2001). The first factor ties into the needs of the individual. For those who promotion focused, their needs are based in growth and development. For those who are prevention focused, their needs are based in promoting a feel of security. Both extremes tie into an individual’s need to avoid pain or to go towards pleasure (Brockner & Higgins, 2001).

Brockner and Higgins (2001) identified the second factor as the types of goals or standards to which an individual attach. Promotion based people “seek to attain goals or
standards associated with the ideal self” (Brockner & Higgins, 2001, p. 38). Prevention people connect to standards or goals that are associated with the self who avoids negative consequences “the ought self” (Brockner & Higgins, 2001, p. 38).

The final factor that creates a person’s regulatory focus is the “psychological situations that matter to people” (Brockner & Higgins, 2001, p. 37). This ties to the comfort that people feel with specific outcomes. With those who are promotion focused, they connect with the presence or absence of positive outcomes. For those who are prevention focused, there is a connection with the presence or absence of negative results. Psychologically, people either are attracted to a specific positive result or attracted to avoiding a specific negative result (Brockner & Higgins, 2001).

In conclusion, RFT is a theory in which the connection people make to their environment depends on whether they are promotion or prevention focused (Brockner & Higgins, 2001). Brockner and Higgins (2001) concluded that when it came to self-regulation, that these two focuses encompassed the mindset people take either to go towards pleasure (promotion focused) or to avoid pain (prevention focused). Managers can have a significant impact on their employees’ self-regulatory processes (Brockner & Higgins, 2001). RFT connects to the CMES because principals’ mindset when looking at goals and standards impacts how they receive feedback while trying to achieve goals or standards. Also, evaluators psychological impact on principals can influence the principals’ overall focus.

**Conclusion: Feedback Intervention Theory and Regulatory Focus Theory**

The purpose of CMES is to provide formative feedback to improve educator performance (State Council for Educator Effectiveness, 2011). With meaningful and
credible feedback to improve practice being the purpose of CMES, the researcher
believes it is helpful to understand a theory of feedback (FIT) and a theory of emotional
connections people make to feedback in their environment (RFT) (Brockner & Higgins,
2001; Kluger & DeNisi, 1996). With the understanding of these theories as a part of this
study’s background, comes an understanding of the importance of paying attention to the
type of feedback principals receive. This especially becomes apparent when this
feedback is to help principals improve their current performance.

**Statement of the Problem**

The stated purpose of the CMES is, “to provide meaningful and credible feedback
that improves performance” (State Council for Educator Effectiveness, 2011, p. 7). This
meaningful and credible feedback is supposed to improve the performance of those being
evaluated (State Council for Educator Effectiveness, 2011). The problem researched was
whether or not the evaluative feedback received by those being evaluated was seen as
meaningful and credible to the participants and if those principals in the sample felt
evaluative feedback improved (or did not improve) performance.

**Research Question**

The researcher has identified the problem as whether or not the feedback received
by those being evaluated was perceived as having value and in turn altered the practice of
those being evaluated. One group of people being evaluated using the CMES is the
principal. The author of this study researched the perception of the CMES system and
the feedback received by principals per the CMES from the principals’ point of view.

The following research questions guided this study:

Q1 What are principals’ perceptions of the Colorado Model Evaluation
System (CMES) evaluation process?
Q2  What are principals’ perceptions of the CMES evaluation feedback?

Q3  According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

**Definition of Key Terms**

*Evaluator* (for a principal) – This the member of a school district or BOCES that is assigned to evaluate a principal. Per the Council of Chief State School Officers (CCSSO) (2015) the role of principal supervisor (or evaluator) falls into three categories:

a. Support and improve principals’ capacity for instructional leadership.

b. Play the role as an effective liaison between the buildings and the district office.

c. Improve their expertise and understanding of their own role as a district leader. (pp. 8-9)

*Feedback* – “Actions taken by (an) external agent(s) to provide information regarding some aspect(s) of one’s task performance” (Kluger & DeNisi, 1996, p. 255).

*Practice* – Acts by principals that relate directly to the performance of the duties of their jobs (State Council for Educator Effectiveness, 2011).

*Principal* – “A person who is employed as the chief executive officer or an assistant chief executive officer of a school in the state who administers, directs, or supervises the education program in the school” (1 CCR 301-87, 2014, p. 2). In this study the assistant principals were asked to participate in the pilot study, and principals were those who were asked to participate in the main study.
Conclusion

The authors of *Race to the Top* required a willingness by states to change their evaluation systems if a state were to win federal funds (US Department of Education, 2009). In Colorado, these changes were initiated by *Executive Order B 2010-001* and continued through *Educator Effectiveness Act*. Both pieces of legislation tasked the State Council for Educator Effectiveness with creating a framework for evaluation (Executive Order B 2010-001, 2010; C.R.S. 22-9-105.5, 2013). While creating this new framework, the State Council for Educator Effectiveness kept five assumptions in mind that guided their work (State Council for Educator Effectiveness, 2011). One of the assumptions used by the Council stated that the main purpose of the new system was to provide feedback to improve performance (State Council for Educator Effectiveness, 2011). To provide a theoretical perspective of feedback, Feedback Intervention Theory and Focus Regulatory Theory were studied (Kluger & DeNisi, 1996; Brockner & Higgins, 2001). In the second chapter, the researcher will provide a comprehensive review of the literature on principals and their evaluations.
CHAPTER II

REVIEW OF LITERATURE

Including the introduction, there are seven sections to this literature review.

Following the introduction, the researcher considers how principal impact has been studied and how long it has taken before a substantive impact was made by a principal. In the third section, the researcher reviews studies that detailed standards, roles, responsibilities, and skills of principals that different researchers believe to be important. The fourth section on principal evaluation was broken into four parts: history of principal evaluation, principal evaluation and feedback studies, researchers’ perceived challenges within evaluation systems; and perceived benefits of the Colorado Model Evaluation System (CMES). In the fifth section the researcher considers the role of the evaluator in the evaluation system. Researchers who used FIT and RFT as a foundation of their work around feedback interventions and peoples’ perceptions of feedback were highlighted in the sixth section (Apodaca et al., 2012; Carlson, Chonko, Kacmar, Neubert & Roberts, 2008; Dijk & Kluger, 2011; Hergovich, Krenn, & Wurth 2013; Khachatryan, 2015; Martin & Mottet, 2011). In the final section the identified gap in the research is discussed.

Mediating Variables: Principal Impact

According to Hattie (2012), teachers have the most significant impact on the academic growth of the students in their classroom. “A student in a high-impact teacher’s classroom has almost a year’s advantage over his or her peers in a lower-effect
teacher’s classroom” (Hattie, 2012, p. 23). School leaders also have an impact on student learning and achievement. School leadership “is second only to teaching among school-related factors in its impact on student learning” (Anderson,, Leithwood, Seashore Louis, & Wahlstrom, 2004, p. 1). A number of studies that identify the mediating variables that help to define principals’ impact on student achievement are discussed below. The initial study summarized by the researcher is Hallinger and Heck’s (1996) meta-analysis in which the authors categorized the different types of effects researchers look for when trying to measure how principals make a difference, and came up with a conclusion based on their research. This study is followed by Coelli and Green’s (2012) analysis on the time it takes for principals’ to make an impact on their schools. After establishing that there have been multiple types of principal effects studied with Hallinger and Heck’s (1996) study and determining the time it takes with Coelli and Green’s (2012) work, the researcher then considered studies that researched what principals’ effects are.

In an empirical review of the literature, Hallinger and Heck (1996) adapted a set of effects models from Pitner to categorize studies of principal effect on student achievement and how the effect occurred according to these studies. These effects models included “Direct Effects Model, Mediated Effects, Antecedent Effects, and Reciprocal-Effects Model” (Hallinger & Heck, 1996, p. 16). These effects models categories were placed in a continuum that labeled how the principal impacts student achievement (Hallinger & Heck). Illustrated in Figure 2.1 is the role of principals based on empirical findings (Hallinger & Heck, 1996).
When comparing studies that fit into the different effects models, Hallinger and Heck (1996) concluded studies that found the principal has an indirect impact on student achievement to be the most accurate. “Although it is theoretically possible that principals do exert some direct effect on students’ learning, the linkage between principal leadership and student learning (as measured by school outcomes) is inextricably tied to the actions of others in the school” (Hallinger & Heck, 1996, p. 24). With this in mind, Hallinger and Heck (1996) also stated that researchers should “focus their attention on uncovering the relationship between principal leadership and those mediating variables that we now believe influence student achievement” (p. 36). After establishing the multiple models that principals have been found to impact student achievement in their schools, the next study that is discussed looked at the time it takes principals to make an impact on student achievement.

Coelli and Green (2012) studied the impact by high school principals on their school environments in a school district in Canada using graduation rates and English
exam scores. They found that there was little to no impact by principals on schools’ performance when testing for immediate impact of principals. Coelli and Green believed that the lagging impact of previous principals could influence the performance of the school and current principal. “It takes time for principals to have their full impact on a school, [and] we find that individual principals can have substantial impacts on both outcomes if given enough time at a school to make their mark” (Coelli & Green, 2012, p. 107). In personal correspondence with David Green, Green distinguished between effect and full effect:

I think you want to distinguish between “an effect” – which might mean a noticeable difference relative to the previous principal - and their “full effect”, i.e., how different student outcomes are once a given principal has had a chance to implement her or his full vision. (Green, personal communication, July, 7, 2014)

For an effect of a new principal to occur, Green also wrote in this personal communication that it takes at least two years. For the full effect of a principal to occur, this takes well over five years (Green, personal communication, July, 7, 2014).

Although they presented multiple different models to assess principal impact, Hallinger and Heck (1996) suggested: “If the impact of principal leadership is achieved through indirect means, we must advance our understanding of how such linkages are shaped by the principal” (p. 34). Marzano, McNulty, and Waters (2005), through a factor analysis connected to 21 leadership responsibilities,² established a link between principals and the schools where they work through first order and second order change.

---
² These 21 responsibilities are described later in this chapter.
First order change is incremental. It can be the next most obvious step to take in a school or a district. Second-order change is anything but incremental. It involves dramatic departures from the expected, both in defining a given problem and in finding a solution. (Marzano et al., 2005, p. 66)

Most people solve issues primarily through first order change (Marzano et al., 2005). “It makes sense that we would tend to approach new problems from the perspective of our experiences-as issues that can be solved using our previous repertoire of solutions” (Marzano et al., 2005, p. 67). For those problems that cannot be solved by organizations’ or schools’ repertoire of experiences, Marzano et al. wrote that these organizations need to build new ways to solve problems. These new ways of thinking are second order change. In schools, “solutions to most recurring modern-day problems require a second-order perspective” (Marzano et al., 2005, p. 68). Ultimately, there is a balance needed by a school principal when working for first order change and that which needs a larger vision and therefore requires second order change.

Ahtaridou et al. (2009) used mixed methods methodology to research principals impact on schools in England and discovered twelve claims that represent all effective school leaders.\(^3\) The principals and schools the researchers selected were chosen because

the schools had students who had significant academic growth over a small period of time (Ahtaridou et al., 2009). Ahtaridou et al. (2009) discovered that these 12 claims proved that there is “an increasingly sophisticated understanding of the strong links that have been traditionally been thought to exist between school leadership and student achievement” (p. 191). The complication of this connection was also highlighted by Ahtaridou et al. (2009) explaining that the findings should not be treated as individual findings, but should be looked at together to “give a deeper insight into both qualities and strategies of leaders that impact positively on pupil progress and achievement” (p. 191). The complicated connection between the 12 claims was made visual by Ahtaridou et al. in a graphic they created that is shown in Figure 2.2. This figure demonstrated, “how the new claims fit together in an integrated way and add value to each other” (Ahtaridou et al., 2009, p. 192). The ultimate impact of principals and their complicated roles are on the achievement of the students in their building.

Distribution of Leadership, 10. Leadership Trust – A Key to Improvement, 11.
Leadership Differences by Improvement Groupings, 12. Leadership Difference by Socio-economic Context. (pp. 183 – 190).
Anderson et al. (2004) completed a report that was the result of their review of the literature, in which they detailed the impact of effective school leaders on student achievement. They first described where the school leaders had the largest effect and ended their report with how the leadership impacts student learning. For the topic on where the leaders have the largest effect, Anderson et al. (2004) came to two conclusions: “1. Leadership is second only to classroom instruction among all school-related factors that contribute to what students learn at school. 2. Leadership effects are usually largest where and when they are needed the most.” (p. 3). Anderson et al. (2004) then asked how the leaders impact student learning and came up with three conclusions: “1. Mostly leaders contribute to student learning indirectly, through their influence on other people or features of their organizations. 2. The evidence provides very good clues about who or what educational leaders should pay the most attention to within their organizations. 3.
We need to know much more about what leaders do to further develop those high-priority parts of their organizations.” (pp. 11-12). School leaders have a strong impact on student learning, and that there was still more research that needed to be done in order to better understand leaders’ impacts (Anderson et al., 2004).

Teachers and principals have impact on student achievement, though the impact manifests itself in different ways (Ahtaridou et al., 2009; Anderson et al., 2004; Hattie, 2012). After establishing how principals’ might be seen to impact student achievement and the time it takes for principals to make an impact (Coelli & Green, 2012; Hallinger & Heck, 1996), three studies were summarized to demonstrate how a principal might make an impact on student achievement (Ahtaridou et al., 2009; Anderson et al., 2004; Marzano et al., 2005). Since Anderson et al. (2004), Ahtaridou et al. (2009), Coelli and Green (2012), Hallinger and Heck (1996), and Marzano et al. (2005) determined that principals do have an impact on student achievement, the next section of this paper explains standards, role conceptions, skills, responsibilities, and character traits needed to best impact principal performance.

**Mediating Variables: Standards, Role Conceptions, Skills, Responsibilities, and Character Traits of Principals**

Hallinger and Heck (1996) identified a need for researchers to “focus greater attention on uncovering the relationship between principal leadership and those mediating variables that we now believe influence student achievement” (p. 36). By identifying these variables, researchers could improve overall school performance by identifying targets for principals to reach. Regarding principal evaluation, “Principal evaluation systems appear to be most effective when they are based upon clear standards and expectations of performance and aligned with the key goals and needs of principals,
schools, and districts” (Davis, Kearney, Sanders, Thomas, & Leon, 2011, p. 35).

According to the following researchers, these mediating variables that best define clear standards and expectations for principal evaluation are found in the roles that principals play, the responsibilities of principals, the skills that principals need to be considered effective, and/or the character traits of strong principals (CCSSO, 2008; Clifford, Behrstock-Sherratt, & Fetters, 2012; Fullan, 2002; Gold, 2003; Hallinger, Wang, & Chen, 2013; Marzano et al., 2005; Matthews & Crow, 2010; Williams, 2008). Roles played, skills needed, responsibilities principals need to demonstrate, and character traits of strong principals are all terms used by the researchers to categorize the things that principals need to excel at their jobs, according to these researchers. Although each of these researchers used different terms to categorize effective principals’ actions, all categorized terms describe what these researchers have found that effective principals do or traits of effective principals.

In Table 2.1 the skills, role conceptions, and responsibilities identified by six researchers and the Interstate School Leadership Licensure Consortium (ISLLC) standards are summarized. The purpose of Table 2.1 is to highlight the similarities and differences within the research.
After creating Table 2.1 the researcher noticed two potential trends. One is the large number of skills identified by the researchers and how many of the researchers

4 The ISLLC Standards were included in this chart because they also have identified a set of skills that are a priority for National Policy Board for Educational Administration (NBPEA).
identified different standards, role conceptions, skills, and responsibilities. The second is that there are some standards, role conceptions, skills, and responsibilities that showed up in almost all the evaluated studies. Together, both trends show that a principal’s job requires a complicated mix of standards, role conceptions, skills, and responsibilities.

Table 2.1 has 24 standards, role conceptions, skills, or responsibilities total. Of those 24, the number of standards, role conceptions, skills, or responsibilities that each researcher identified as most important to the role of principal varies greatly. The largest number identified by one researcher or one set of researchers is the 20 identified by Marzano et al. (2005). Gold (2003) identified four skills. The other five researchers or organizations identified a range of eight to 13. These numbers demonstrate that a large variety of researchers and organizations have identified a large variety of necessary standards, role conceptions, skills, or responsibilities of principals. The number of researchers and organizations that identified different necessary standards, role conceptions, skills, or responsibilities also shows that the complicated job of the principalship has a variety of interpretations of effectiveness.

There are three role conceptions, skills, or responsibilities that were identified by a large number of the researchers. These three are: intellectual stimulation, implementation and knowledge of curriculum, instruction, and assessment, and monitoring and evaluating. CCSSO (2008), Fullan (2002), Gold (2003), Hallinger et al. (2013), Marzano et al. (2005), and Matthews and Crow (2010) all identified intellectual stimulation as a critical skill of a principal. Marzano et al. (2005) defined intellectual stimulation as “ensures faculty and staff are aware of the most current theories and practices and makes the discussion of these a regular aspect of a school’s culture” (pp. 42
Implementation and knowledge of curriculum, instruction and assessment is a high priority skill for principals per CCSSO (2008), Clifford et al. (2012), Fullan (2002), Hallinger et al. (2013), Marzano et al. (2005), Matthews and Crow (2010). Marzano et al. (2005) had this skill split between knowledge of and implementation of, but defined both as the principal “is knowledgeable about” and “is directly involved in the design and implementation of curriculum, instruction, and assessment practices” (pp. 42 – 43). The third skill that was included by five out of the six researchers and the ISLLC standards was monitoring and evaluating. Marzano et al. (2005) defined monitoring and evaluating as when a principal “monitors the effectiveness of school practices and their impact on student learning” (pp. 42 – 43). All three of these skills/responsibilities highlight the importance of principals supporting classroom instruction.

The first piece of research detailed and described below will be the ISLLC standards. It is its own separate category, because these standards were (and continue to be) used by universities in their principal development programs and states for their principal evaluation systems (Canole & Young, 2013). These standards were also the standards referenced while creating the CMES system in Colorado. In the section identified as: Role Conceptions, Skills, Responsibilities and Character Traits of Principals the seven researchers show the perspectives of these researchers on what makes an effective principal.

**Interstate School Leaders Licensure Consortium Standards**

In the 1990s, standards and the definition of the principalship became a priority of many states, educational leaders, and universities (Murphy, 2003). For this reason, a group of experts came together to create standards to be used by states and universities.
The *ISLLC Standards* were created by the National Policy Board for Educational Administration (NPBEA) to “take up the challenging task of designing the first set of national standards for educational leaders” (Canole & Young, 2013, p. 5). The final product of their work was the *Interstate School Leaders Licensure Consortium Standards for School Leaders*. The NPBEA adopted and released these standards in 1996. As of 2013, 41 states were using these standards in some capacity (Canole & Young, 2013). By 2014, Vogel and Weiler (2014) noted all states were at least referencing the *ISLLC Standards*.

By creating and implementing these standards, the consortium hoped to, “provide the means to shift the metric of school administration from management to educational leadership and from administration to learning” (Murphy, 2003, pp. 16-17). Through the creation of these standards and the editing thereof in 2008, the consortium was also hoping to take this proposed shift in educational leadership and ensure its impact through the creation and adoption of these standards.

The following Standards are the Educational Leadership Policy Standards: *ISLLC 2008* “as adopted by the National Policy Board for Educational Administration on December 12, 2007” (CCSSO, 2008):

- Standard 1: An educational leader promotes the success of every student by facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by all stakeholders.
- Standard 2: An education leader promotes the success of every student by advocating, nurturing, and sustaining a school culture and instructional program conducive to student learning and staff professional growth.
- Standard 3: An education leader promotes the success of every student by ensuring management of the organization, operation, and resources for a safe, efficient, and effective learning environment.
- Standard 4: An education leader promotes the success of every student by collaborating with faculty and community members, responding to diverse community interests and needs, and mobilizing community resources.
- Standard 5: An education leader promotes the success of every student by acting with integrity, fairness, and in an ethical manner.
- Standard 6: An education leader promotes the success of every student by understanding, responding to, and influencing the political, social, economic, legal, and cultural context. (CCSSO, 2008)

The NPBEA felt that the ISLLC Standards, both in their original form in 1996 and in their revised form in 2008, were important for states and local educational agencies as they developed their own standards, universities and colleges as they continue to refine their Educational Leadership programs, and educational leaders as they develop and refine their skills as leaders (CCSSO, 2008).

In the state of Colorado, the NPBEA’s intended shift of principal leadership from manager to instructional leader has occurred. Colorado was one of states that used these standards as a reference point when making the Colorado Model Evaluation System.
For CMES, the *ISLLC Standards* provide a research-based framework on which to base an evaluation system.

**Role Conceptions, Skills, and Responsibilities of Principals**

Although the *ISLLC Standards* are used by states to develop evaluation systems, there are multiple researchers who identified role conceptions, skills, or responsibilities they believed to be indicative of effective principals (Clifford et al., 2012; Fullan, 2002; Gold, 2003; Hallinger et al., 2013; Marzano et al., 2005; Matthews & Crow, 2010; Williams, 2008). The role conceptions, skills, and responsibilities were identified in Table 2.1. Below is a short explanation of each of the researchers’ perspectives that were highlighted in Table 2.1. The role conceptions, skills, and responsibilities emphasized by these researchers show the impact of the principal on student achievement is complicated and indirect but important.

In 2005, Marzano et al. (2005) published a meta-analysis of leadership behaviors of school leaders. In this meta-analysis, Marzano et al. identified and described 21 responsibilities of school leaders. The authors acknowledged that, “to a great extent, our findings validate the opinions expressed by leadership theorists for decades” (Marzano et al., 2005, p. 41). They drew the strongest comparison of their research to Cotton’s (2003) 25 responsibilities identified in *Principals and Student Achievement: What the Research Says*. However, Marzano et al. (2005) believed that their “21 responsibilities provide some new insights into the nature of school leadership” (p. 41). The 21 responsibilities of school leaders, as identified by Marzano et al. (2005) are: affirmation, change agent, contingent rewards, communication, culture, discipline, flexibility, focus, ideals/beliefs, input, intellectual stimulation, involvement in curriculum, instruction, and assessment,
knowledge of curriculum, instruction, and assessment, monitoring/evaluating, optimizer, order, outreach, relationships, resources, situational awareness, and visibility.

Matthews and Crow (2010) identified eight role conceptions that they believed best defined the role of principal and assistant principal. The authors believed that the three roles of learner, culture builder, and advocate are “foundational to the principal’s primary role conception of leader” (Matthews & Crow, 2010, p. 14). The final five role conceptions “contribute to the teaching and learning in a professional learning community: leader, mentor, supervisor, manager, and politician” (Matthews & Crow, 2010, p. 14). The authors further defined these role conceptions and explained whether the principals’ responsibilities in each of these roles are facilitative or directive. They then categorized the impact on each of these conceptions as either direct or indirect. Whether direct or indirect, Matthews and Crow (2010) believed these role conceptions to be integral to the overall role of school principal.

Hallinger et al. (2013) conducted a meta-analysis of the research that studied a tool created by Phillip Hallinger called the Principal Instructional Management Rating Scale (PIMRS). The PIMRS outlines the skills of strong principals. These skills of an effective principal were divided into three dimensions. The first dimension is: defining the school mission. Under defining the school mission, the subcategories are (1) frames the school’s goals and (2) communicates the school’s goals. The second dimension was: managing the instructional program. Under this dimension, Hallinger’s sub-categories are (1) coordinates the curriculum, (2) supervises and evaluates instruction, and (3) monitors student progress. The final dimension of the PIMRS is: developing the school-learning climate. The subcategories of this dimension are: (1) protects instructional time,
(2) provides incentives for teachers, (3) provides incentives for learning, (4) promotes professional development, and (5) maintains high visibility. Researchers and practitioners, using the PIMRS tool, evaluate the instructional leadership capabilities of principals (Hallinger et al., 2013).

In 2002, Michael Fullan wrote an article highlighting the need for change in priorities in education from standards to strong leadership. “System transformation of the type educators now aspire to cannot be accomplished without first ensuring solid leadership at all levels of the system” (Fullan, 2002, p. 14). Acknowledging the complicated role of leaders in education, Fullan wrote that the changes and expectations of educational leaders required research beyond current literature at that time. He highlighted the duality of an ideal school system. “You need great conditions to develop the kind of leaders we have been talking about, and you need great leaders to develop the conditions that will produce great leaders” (Fullan, 2002, p. 16). Fullan (2002) suggested five interrelated themes that “enhance the chances of sustainability” for the leaders of schools (p. 16). These five themes are: “opportunity and depth of learning, policies for individual development, learning in context and systemness, leadership succession and leaders at many levels, and improving the teaching profession” (Fullan, 2002, pp. 16-17). By focusing on these five themes, Fullan (2002) believed leaders improve their chances for success.

Gold (2003) wrote about the concepts of transformational and instructional leadership within the role of principal. Transformational leadership “focuses on the people involved – relationships between them, in particular – and requires an approach that seeks to transform staff feelings, attitudes and beliefs” (Gold, 2003, p. 128).
Instructional leadership “typically assumes that the critical focus for attention by school leaders should be the behaviors of staff as they engage in activities directly affecting the quality of teaching and learning in pursuit of enhanced pupil outcomes” (Gold, 2003, p. 128). With the mindset that transformational leadership and instructional leadership should be the focus of successful principals, Gold determined themes that demonstrated leadership in action based on 10 case studies of principals in England who were considered exceptional. Four general connections were made: working with, managing, and even searching out change; paying careful attention to information management within the school – thus keeping staff constantly informed; working very closely and sometimes seamlessly with their leadership groups; and developing leadership capacity and responsibility throughout their schools (Gold, 2003). Within these four themes lay a balance between transformational and instructional leadership. Great principals considered the balance between being a transformational leader and an instructional leader of highest importance to succeed with “the wider educational, social and personal development of all pupils and staff” (Gold, 2003, p. 136).

Clifford et al. (2012) found the changing roles of leadership require the principal to change roles as a leader to address specific needs of schools. “While different visions of school leadership are apparent, we note that a single leader will assume multiple leadership roles depending on the leadership context” (Clifford et al., 2012, p. 2). These roles are traditional manager, supervisor of standards, adaptive leader, instructional leader, and leader among leaders. Clifford et al. believed principals’ roles in a school have become increasingly complicated as they indicated in their five roles played by all
principals. Even with complicated roles, the expectation of effectiveness is maintained no matter the role being played (Clifford et al., 2012).

**Character Traits of Principals**

Williams’ (2008) work was not included in Table 2.1 because the competencies she researched related more to character traits than to competencies, skills, role conceptions researched by the other authors. The researcher of this dissertation felt Williams’ research needed to be included because mediating variables to perform the job of principal, may include who the principal is not just what they need to do. Williams stated the character traits of outstanding urban school principals are based in both social and emotional intelligence competencies. “Emotional Intelligence [EI] competencies contribute to an individual’s ability to understand and manage oneself. Social intelligence competencies focus on an individual’s ability to understand others and manage relationships” (Williams, 2008, p. 39). Based in social and emotional intelligence, Williams identified six competencies for outstanding urban principals: self-confidence, achievement orientation, initiative, organizational awareness, leadership, and teamwork/collaboration. Those principals who were more skilled in both emotional and social intelligences could handle more complex and dynamic. Being strong in these six critical competencies indicated the potential ability to succeed in a complex environment therefore supporting the achievement of students and staff. These six character traits identified how principals need to function based on two intelligences that help them relate to people personally and individually (emotional intelligence) and in groups (social intelligence) (Williams, 2008).
Conclusion: Role Conceptions, Skills, Responsibilities, and Character Traits of Principals

Though there are many similarities between the researchers’ conclusions to the most important role conceptions, skills, responsibilities, and character traits there are many different determinations made about that which principals need to do to impact student achievement. Matthews and Crow’s (2010) book was used in one of the researcher’s doctoral classes, and provided a starting point for this section of this Literature Review. Two of the researchers are considered experts in the field of principal leadership and the skills they identified would be well respected in the field (Fullan, 2002; Marzano et al., 2005). Though the role of principal is a complicated one, the PIMRS analyzed by Hallinger et al. (2013) helped to provide a framework for evaluating the roles that might be used by researchers and is not part of a principal’s evaluation. Helping to be a leader of change in a building and the human connection and understanding needed to make these changes are the reasons for the use of Gold (2003) and Helen Williams (2008). Gold (2003) tied Transformational Leadership with Instructional Leadership, which connects the need for principals to transform staffs to be better at instructing students and this article has been cited 230 times on Google Scholar. Williams’ (2008) connecting social and emotional intelligences to the principalship demonstrates how the root of a successful principalship is how a principal connects with people. The final brief written by Clifford et al. (2012) was published after RttT would have been implemented and would provide a view of skills needed by today’s principals. Overall, these researchers provided a good foundation for a review of principals’ skills, role-conceptions, responsibilities, and character traits (Clifford et al., 2012; Fullan, 2002; Gold, 2003; Hallinger et al., 2013; Marzano et al., 2005; Matthews & Crow, 2010;
Williams, 2008). This foundation was needed when studying the making of an effective principal evaluation in order to grow or support effective principals.

**Principal Evaluations**

To ensure that schools are successful, the education system must put “resources into recruiting, hiring, developing, and keeping the best teachers and leaders” (Stronge, 2013, p. 5). Since *Race to the Top*, evaluation has been identified as a critical tool to ensure the best educators are working with our students (C.R.S. 22-9-105.5, 2013; Executive Order B 2010-001, 2010; US Department of Education, 2009). Stronge (2013) concluded that “Unless we have effective evaluation systems in place that accurately differentiate performance, we simply can’t even discern whether or not we have effective principals” (p. 5). Evaluations should also help with principal professional development.

If we hope to improve principals’ performance, we must be able and willing to provide honest, accurate, and meaningful evaluation feedback in order to identify areas in need of improvement and enable principals to make informed decisions regarding professional development to bridge the gap between current practices and desired performance. (Stronge, 2013, p. 6)

A strong evaluation system with effective evaluators can provide both an understanding of principal performance and feedback to improve principal performance.

The following sections will summarize broad topics of principal evaluation starting with the history of principal evaluation in the United States. The section will then continue to narrow the focus by looking a Colorado’s reaction to *RttT* through the recommendations created by the State Council for Educator Effectiveness (2011). The next focus of the principal evaluation section will be a summary of the process of
evaluation for educators created by the Colorado Department of Education. The conclusion will be a report of the results of the first year of the CMES for principals (2014-2015).

**History of Principal Evaluation**

Principal evaluation has been practiced by school districts since the early 20th Century (Ginsberg & Berry, 1990). “As the ideas of scientific management and efficiency of operation spread throughout the school systems in the country, means of rating principals were devised and discussed in educational journals” (Ginsberg & Berry, 1990, P. 207). In the late 1960s, a movement for school accountability developed from “widespread dissatisfaction with the quality of education students received in their schools” (Ginsberg & Thompson, 1992, p. 65). From the 1960s to the late 1970s the push of this movement was on the increased responsibility of student learning on the teachers and administrators through evaluation and other measures (Ginsberg & Thompson, 1992). By 1984, the number of states that mandated formal principal evaluations numbered 27 and the number of school districts reporting formal principal evaluations were up to 85.9 percent (Snyder & Ebmeier, 1992). By a 2009 study authored by Goldring et al. (2009), “virtually every school district in the United States requires some form of evaluation of its principals” (p. 20).

Over the last 30 years there has been a transition in expectations of principals, which has impacted evaluations. With the creation of the *ISLLC Standards* and the larger accountability systems to which they relate has come an increase in results rather than inputs when it comes to principal assessments (Portin, Feldman, & Knapp, 2006). In *Reculturing Educational Leadership: The ISLLC Standards Ten Years Out* (Murphy,
Joseph Murphy, one of the creators of the ISLLC Standards, wrote of the transition of focus of effective schools and their administrators. He wrote that effective schools were transitioning from teachers and their delivery of instruction to what and how students were learning. He also stated that effective schools were transitioning from hierarchical and bureaucratic decision-making structures to flattened school structures whose purposes were to support learning. According to Murphy, the final perceived change in school structures were schools being communities as opposed to organizations. Overall, the impact of these changes highlighted the transition of school leaders to instructional leaders which was at the core of the creation of the ISLLC Standards (Murphy, 2003). With all the states at least referencing the ISLLC Standards currently (Vogel & Weiler, 2014), the ISLLC Standards have the potential to impact all principals in the United States.

**Race to the Top.** The basics of *Race to the Top (RttT)* were outlined in Chapter I, as was the legislation for Colorado specifically based on Colorado’s application for RttT. This section will give the background, history, and the intended philosophy of RttT. The initial paragraph starts with the number of states that were impacted in some capacity by RttT.

Over the two rounds of applications for RttT applications were submitted by all but four states, with 40 states applying for the initial round (McGuinn, 2011). In the initial application process “15 of them [states] (including some strong union states like California, Michigan, and Ohio) passed revised regulations in advance of their applications to improve their chances of winning” (McGuinn, 2011, p. 143). The Obama
administration claimed that significant shifts occurred in education policies in 34 states due to *RttT* (McGuinn, 2011).

According to the original director of the *RttT* program (Joanne Weiss)\(^5\) the philosophy behind the creation of *RttT* had three parts:

- Shifting the federal role from a focus on means to a focus on ends, shifting from sanctions to incentives as a way of motivating state reform, and shifting the Department of Education away from being a compliance-monitoring organization to being one focused on capacity building and innovation. (McGuinn, 2011, p. 140).

McGuinn also stated that the main policy shifts where this impact was proposed to occur was in the acceptance and promotion of charter schools as a priority in the grant application process, the shift in accountability measures for teachers in their evaluations, and the proposed creation of common standards and common assessments by states as a part of the application process. McGuinn (2011) also concluded that putting Governors and State Education Executives in charge of the application process changing their role from rubber stamping education in their state to making education a part of the states’ political dialogue was an attempt by the Obama administration to create meaningful and long-lasting change in education.

Regardless of intended outcomes of *RttT* many states made changes to their laws to win the competitive grant (McGuinn, 2011). In Colorado, those changes occurred through *Executive Order B 2010 - 001* and *The Educator Effectiveness Act*. One big change the writers of the *Educator Effectiveness Act* wanted was a wholesale change to

---

\(^{5}\) As cited in McGuinn (2011)
educator evaluation systems in the state of Colorado. The following sections detail the proposed outline for the evaluations by the State Council for Educator Effectiveness in the *State Council for Educator Effectiveness: Report and Recommendations (2011)* and further details the evaluation system created for educators by CDE.

**Effort in Colorado to change educator evaluation.** The State Council for Educator Effectiveness’ (the Council) first defined an effective principal, then explained the purpose of the new evaluation system, and finally detailed how the evaluation system should be developed. The final product of the *State Council for Educator Effectiveness: Reports and Recommendations* was a skeleton on which an evaluation system was built, and supposed to be continuously improved. The hope was to create and support effective educators in Colorado’s schools (State Council for Educator Effectiveness, 2011).

The Council was to first define what effective principal means (C.R.S. 22-9-105, 2013; Executive Order B 2010-001, 2010; State Council for Educator Effectiveness, 2011). The Council created the following definition:

*Effective principals in the state of Colorado are responsible for the collective success of their schools, including the learning, growth and achievement of both students and staff. As the school’s primary instructional leader, effective principals enable critical discourse and data-driven reflection about curriculum, assessment, instruction, and student progress, and create structures to facilitate improvement. Effective principals are adept at creating systems that maximize the utilization of resources and human capital, foster collaboration, and facilitate*

---

6 The original legislation was in 2011. The version of this legislation used in this paper was edited and completed in 2013.
constructive change. By creating a common vision and articulating shared values, effective principals lead and manage their schools in a manner that supports the school’s ability to promote equity and to continually improve its positive impact on students and families. (State Council for Educator Effectiveness, 2011, pp. 14 – 15)

The Council stated effective principals are responsible for collective success of, instructional leadership of, system creation for, and mission and vision building for the whole school (State Council for Educator Effectiveness, 2011). This responsibility for the whole school is highlighted later in the report as well. “The principal framework emphasizes the fact that principals are responsible for the success of their school overall, including the success of all students and all teachers in the school” (State Council for Educator Effectiveness, 2011, p. 94). The Council identified principals as responsible for the collective success of the school.

The Council decided on four purposes of evaluation, as it was defined in the Educator Effectiveness Act. These four purposes were as follows:

Serve as a basis for the improvement of instruction; enhance the implementation of programs of curriculum; serve as a measurement of the professional growth and development of licensed personnel; and provide a basis for making decisions in the areas of hiring, compensation, promotion, assignment, professional development, earning and retaining non-probationary status, dismissal, and nonrenewal of contract. (State Council for Educator Effectiveness, 2011, p. 32).
With the four purposes, the Council believed that the Educator Effectiveness Act changed how performance was determined from being based on inputs to being based on results (State Council for Educator Effectiveness, 2011).

After clarifying the purposes of the Council, the Council then explained its recommendations (State Council for Educator Effectiveness, 2011). To help clarify the recommendations, the Council created a graphic to show the state framework for principal evaluation. This framework was designed as the foundation or skeleton that every school district would follow no matter if a school district plans on using the full state developed evaluation or created a district-specific evaluation. The creators of this graphic demonstrated the connection of all parts of the system to one another to produce a final rating of a principal evaluation as ineffective, partially effective, effective, or highly effective (State Council for Educator Effectiveness, 2011).

The Council wrote about the use of feedback throughout the report and recommendations (State Council for Educator Effectiveness, 2011). “Today’s educators need and deserve an evaluation process that is designed to drive substantive feedback and vigorous conversations about effective teaching and effective school leadership” (State Council for Educator Effectiveness, 2011, p. 32). Frequent feedback must also be provided to principals. “Districts shall collect evidence of principal performance with enough frequency to ensure that principals are provided with ongoing feedback and the opportunity to improve performance” (State Council for Educator Effectiveness, 2011, p. 109). Since there is a need for continuous feedback to improve, “data collection, feedback, and opportunity for improvement should be structured to create an ongoing

---

7 The graphic of the framework is shown in Appendix A.
evaluation process rather than an annual event” (State Council for Educator Effectiveness, 2011, p. 108). Feedback needs to be both substantive and frequent to ensure success of the new Colorado Model Evaluation System.

In Colorado, under the new principal evaluation system, school districts or Boards of Cooperative Education Services (BOCES) must use, “the state framework, the state definition for effective principals, and the quality standards for principals, and must use multiple measures to assess performance that include data about teacher and staff perceptions and the School Performance Framework” (State Council for Educator Effectiveness, 2011, p. 95). Within these multiple measures, the Council allowed for some flexibility for school districts and required the principals who are to be evaluated to have some representation while determining the other measures of performance. “Districts shall involve principals in the district, including members of the representative association if one exists, in developing or adopting tools to measure a principal’s performance of the Principal Quality Standards” (State Council for Educator Effectiveness, 2011, p. 108).

The school district, with input from this representative principal group, then should determine the measures of performance and how all measures will be weighted in the overall evaluation (State Council for Educator Effectiveness, 2011). However, this weighting is restricted. “Standards I – VI determine no more than 50% of the principal’s performance; and the measures of Standard VII (student growth) determine at least 50% of the weight of the evaluation” (State Council for Educator Effectiveness, 2011, p. 108).
Student performance and growth has the heaviest weight compared to other standards in the evaluation.\textsuperscript{8}

The Council adopted seven quality standards that were developed, prior to the \textit{Educator Effectiveness Act}, by the School Leadership Academy Board (SLAB) (State Council for Educator Effectiveness, 2011). These seven quality standards focus on all the different roles the principal plays. The Standards include:

- Standard I – Principals demonstrate strategic leadership,
- Standard II – Principals demonstrate instructional leadership,
- Standard III – Principals demonstrate school culture and equity leadership,
- Standard IV – Principals demonstrate human resources leadership,
- Standard V – Principals demonstrate managerial leadership,
- Standard VI – Principals demonstrate external development leadership,
- Standard VII – Principals demonstrate leadership around student growth.

(State Council for Educator Effectiveness, 2011, pp. 15-17)

These standards are the basis of the evaluation, and the elements under each of the standards show the expected behaviors and tasks a principal must accomplish to receive a score that demonstrates levels of performance (State Council for Educator Effectiveness, 2011, pp. 15-17).

\textsuperscript{8} For the 2014-15 school year only, Senate Bill 14-165 provides school districts and BOCES additional flexibility with respect to how they rate educators on measures of student learning (Quality Standard VI for teachers and VII for principals). For a single school year, districts and BOCES may weight the measures of student learning/outcomes rating anywhere between zero and 50 percent (Colorado Department of Education, August 2014, p. 104).
They are based in the *Interstate School Leaders Licensure Consortium (ISLLC)* standards created in the 1990s and revised in the late 2000s.\(^9\)

Appendix B is a sample state-scoring framework for the Principal Evaluation (State Council for Educator Effectiveness, 2011). This score is a combination of the score that a principal receives from his evaluators on the Quality Standards I-VI and the student growth score (Standard VII) from the group attribution of the school. This final score will then label the principal ineffective, partially effective, effective, or highly effective. (State Council for Educator Effectiveness, 2011).

Principals are to develop professional performance plans with their evaluators (State Council for Educator Effectiveness, 2011). The requirements of the professional performance plans are three-fold: First, principal performance plans must have goals; Second, principals must have an explicit plan in place to address the numbers of effective teachers in their buildings; Finally, when developing goals, the plan shall include goals that address school climate:

- Developed with reference to the biannual Teaching, Empowering, Leading, and Learning (TELL) initiative survey, when available, and other appropriate data, including conditions highlighted in Comprehensive Appraisal for District Improvement and School Support Team reviews. (State Council for Educator Effectiveness, 2011, p. 109)

These three requirements help to prioritize what the Council believed to be of highest importance for principals to best support student achievement.

\(^9\) These standards can also be known as the Educational Leadership Consortium Council (ELCC) Standards 2008 (Vogel & Weiler, 2014).
The Council created a comprehensive and structured skeleton on which an evaluation is to be built (State Council for Educator Effectiveness, 2011). This system is an attempt by the Council to categorize the complexity of a principal’s job, define effective principals, and provide principals a system that supports “the principal’s growth and development while simultaneously holding him or her accountable for student success” (Stronge, 2013, p. 8).

**Colorado Model Evaluation System.** The following paragraphs state the details of the system as it was created by CDE. The information was gathered from *The User's Guide: Colorado State Model Educator Evaluation System: 2016-2017* (Colorado Department of Education, 2016-2017). The steps of the CMES include:

![Visual Model of CMES process for all educators](image-url)

*Figure 2.3. Visual Model of CMES process for all educators (Colorado Department of Education, 2016-2017, p. 16)*

To follow it with fidelity, the evaluators and principals must follow the nine parts of the evaluation sequentially (Colorado Department of Education, 2016-2017). The different steps explained in *The User’s Guide* are summarized in the following sentences.

Training and annual orientation are the two parts of the evaluation that were created to
familiarize the evaluator and the person being evaluated to the system. Training is to be done prior to using the CMES. This training must be done “to ensure that everyone has the foundational knowledge needed to implement the system” (Colorado Department of Education, 2016-2017, p. 17). Staff members also need to go through an orientation to see that which is new to the system and changes to the process. After the orientation, and by the end of the first month, the principal being evaluated will need to complete the self-assessment. During principals’ self-assessment they would look through the principal rubric created by CDE, assess where they see themselves scoring well, and where their perceived deficiencies might be. Also, to be done by the end of the first month is the professional growth plan. This professional growth plan will include goals that are to be set by the principal being evaluated. These goals are to be compared to district and school initiatives and goals to ensure alignment. This plan needs to be signed off on by both the evaluator and the principal being evaluated. Prior to the second semester of the school year, the evaluator and principal must have a meeting in which the goals that were agreed upon by both the evaluator and principal. This meeting is called the mid-year review. During this meeting “they discuss barriers to completing goals and refine existing goals as needed” (Colorado Department of Education, 2016-2017, p. 17). The sixth step in CMES process is to be completed prior to the final evaluation meeting. This step requires the evaluator to complete the rubric assessing the performance of the principal. After completing the rubric, and three weeks prior to the end of the school year, a meeting must be had between the evaluator and the principal being evaluated.

10 The Colorado Principal and Assistant Principal Rubric was retrieved from: https://www.cde.state.co.us/educatoreffectiveness/rubric-for-colorado-principals (Colorado Department of Education (2014-2015)}
During this meeting the principal and the evaluator “should discuss professional practice ratings and measures of student learning/outcomes, artifacts and any other evidence needed to confirm the accuracy of ratings” (Colorado Department of Education, 2016-2017, p. 17). If they come out of this meeting with an agreement, a final effectiveness rating is given. If they do not agree they then need to meet again to discuss additional information. The ninth and final step in the evaluation process is to set forth how the principal being evaluated will look to improving into the next year. The following year’s professional growth plan could be set during this ninth step (Colorado Department of Education, 2016-2017). Feedback is a priority during this evaluation process, and where it is prioritized in the evaluation cycle is highlighted below.

As one of the five priorities of the evaluation cycle, feedback is important throughout all nine steps of this process (Colorado Department of Education, 2016-2017). The authors of the User’s Guide: Colorado State Model for Educator Effectiveness direct evaluators after the mid-year review to “provide ongoing feedback based on multiple school visits, data, targeted development activities and other information” (Colorado Department of Education, 2016-2017, p. 21). For the principal, these sources of feedback are also supposed to include student, parent and teacher perceptions of the principal’s performance (Colorado Department of Education, 2016-2017). Feedback was one of the five priorities of the evaluation process (Colorado Department of Education, 2016-2017; State Council for Educator Effectiveness, 2011), and the importance of feedback is mentioned often throughout both the State Council for Educator Effectiveness: Report and Recommendations (2011) and The User’s Guide: Colorado State Model for Educator Effectiveness (Colorado Department of Education, 2016-2017).
2014 – 2015 principal final effectiveness ratings data. Though the purpose of the CMES is to provide formative feedback, the result of the CMES process is a final determination of effectiveness (Final Effectiveness Rating) through a calculation of the professional practices with student achievement and growth data (State Council for Educator Effectiveness, 2011; Colorado Department of Education, 2016-2017). The Colorado Department of Education provides access to effectiveness ratings for teachers and principals of the evaluation through the School View Data Center (Colorado Department of Education, 1999-2013). The data detailed below was for the 2014-2015 school year and gathered from the School View Data Center. The percentage of the principals to receive specific effectiveness ratings were: 30 percent Highly Effective, 54 percent Effective, 5 percent Partially Effective, and 11 percent Not Rated. School View did not show the data for ineffective because there were too few to be included in the calculation. School View also provides a breakdown of ratings for novice principals (principals in their first three years) compared with experienced principals (four or more years). Of the total number of novice principals, 24.39 percent were rated Highly Effective, 55.63 percent were rated Effective, 4.23 percent were Partially Effective, and 7.04 percent were Not Rated. The N value for ineffective principals was less than five. For experienced principals 41.47 percent Highly Effective, 49.07 percent Effective, 2.67 percent Partially Effective, and 6.4 percent Not Rated. Significantly more experienced principals were rated Highly Effective (17.08 percent) verses novice principals. There was only a 3.6 percent difference between novice and experience in the Partially

11 The researcher of this paper found that School View included information for all the school districts in the state of Colorado, not just the school districts that use the CMES. (ex. Denver 1 was included)
Effective Rating (Colorado Department of Education, 1999-2013). Overall, no matter the experience of the principal, very few principals received evaluations that placed them in the less than effective category with 95 percent of the principals being rated Effective or Highly Effective.

The CMES system was outlined by The Council, guided by the State Board of Education, and built by CDE. In the 2014 – 2015 school year, ninety-five percent of the principals received at least an Effective rating as the result in the system (Colorado Department of Education, 1999-2013). The researcher of this dissertation attempted to find the perspective of all of these principals on whether or not they believe the feedback according to this system has impacted their practice.

**Studies on Principal Evaluations**

The research of this study will be based in perceptions by principals of the feedback they have received using the Colorado Model Evaluation System. The following is a look at research done on principals’ perceptions of their evaluations. First, the works of four researchers whose conclusions were based on principals’ perceptions are detailed. Next comes a review of a literature review. After the literature review comes the conclusions of researchers who reviewed the principal evaluations of all fifty states and Washington D.C. after RttT was created.

The following study would have occurred after the first round of ISLLC standards and before RttT. In a dissertation on feedback received by principals from their supervisors, Johnson (2005) highlighted four themes: “(a) frequency of feedback, (b) multiple sources of data, (c) types of feedback, (d) enhancing professional growth” (p. 106). The frequency of feedback varied from school district to school district, as did the
time that the principals received individual feedback from their supervisors. The multiple sources of data included: “(a) observations and interactions with the elementary principals or on-site visits, (b) staff surveys, (c) parent surveys and telephone interviews, (d) written reflections by elementary principals’ on accomplishments of professional and personal self-development goals, and (e) superintendents’ annual written summaries” (Johnson, 2005, p. 109).

The principals in this study received formative and summative feedback in different ways (Johnson, 2005). Some received verbal feedback and then written feedback at the end of the year. Others received just the end of the year written feedback. One received written feedback the first year and then verbal feedback for the following years (Johnson, 2005).

The final finding of Johnson (2005) for feedback in relation to the principals’ evaluations was professional growth. All but one of the 21 principals believed that feedback was a necessary part of their professional growth (Johnson, 2005). These findings led Johnson (2005) to the following conclusions about feedback:

- Principals believe that goal setting, feedback, and consistency contribute to the fairness of evaluation systems.

- Frequency of feedback varies between supervisors and elementary principals.

- Multiple sources of data provide information for [evaluators] to give formative and summative feedback throughout the school year.

- Formative and summative feedback facilitates elementary principals’ professional growth.
• [Evaluators] encourage elementary principals to engage in reflective practice, but self-evaluation was not part of Southeast Nebraska evaluation systems. (pp. 133 – 136)

Many of these points are points of emphasis in the CMES.

In a dissertation by Rouse (2009), the author asked questions on the satisfaction of feedback to principals in Massachusetts as the feedback is related to performance standards or school districts’ or principals’ goals. Only 41 percent of the public high school principals in this study were satisfied with the amount of feedback received (Rouse, 2009). However, when the participating principals did receive feedback specific to performance standards or goals, this feedback was viewed as helpful. Rouse (2009) found that “principals who receive help in obtaining useful feedback are more likely to believe that such feedback is of excellent quality and relates to many or all of the goals for which they are held accountable” (p. 108). Informal feedback received did not receive as much credibility from principals. “Informal feedback may be useful to some principals, but a majority of those surveyed believes informal feedback has limited usefulness because it does not relate to the goals and performance standards for which they are being held accountable” (Rouse, 2009, p. 109). The amount of feedback received was a big concern for principals. “Principals in Massachusetts are receiving less than a sufficient amount of feedback to help them attain goals and performance standards” (Rouse, 2009, p. 107).

The third study examined principals’ perceptions of the principal evaluation system in Minnesota (Muench, 2014). Muench (2014) researched three questions: What were principals’ perceptions of their past evaluation practices? Which job duties
highlighted by the evaluation do they spend the most time on? And, What are the perceptions of principals of student assessment data being used on evaluations? The survey was sent to all 582 secondary school principals in Minnesota with 124 responding to the survey (21.3 percent) (Muench, 2014). When responding to past evaluation practices, 90.2 percent of Minnesota secondary principal respondents felt past evaluations were fair. However, “only half of the respondents reported that past evaluations have been valuable or highly valuable for their professional growth” (Muench, 2014, p. 297).

Muench asked about job duties that were highlighted by the evaluation system and should be prioritized, and instructional leadership was the highest priority followed by acting with integrity and ethics and creating a vision. When looking at where principal’s time is prioritized, Muench found that managing the building was the highest rated job duty.

Finally, when asked what their evaluators stressed with the evaluation, Muench found that instructional leadership again was the highest priority for this group of principals. When asked about the role of student achievement data in the evaluation, these principal respondents believed that student assessment data should be a part of the evaluation, “just not at a rate of 35%” of the total for the final evaluation (Muench, 2014, p. 298).

Parylo, Zepeda, and Bengtson (2012) conducted a phenomenological study about principals’ perspectives of their own evaluation. They interviewed sixteen principals from four school districts in Georgia. Eight central themes came from their investigation: increased awareness; process, not an event; transparency; dialogue; trust and respect; feedback; support, and tensions. Increased awareness had to do with the awareness of the principals that they are being evaluated and are “more aware of the strings attached to their evaluation” (Parylo et al., 2012, p. 223). When the theme of “process, not an event”
was explained by the researchers, they compared the current evaluations to how evaluations used to be. The principals stated the old evaluation was once a year and was an event; whereas, the current evaluation system is something that occurs every day, especially in their instructional leadership role (Parylo et al., 2012).

Transparency was extremely important to these principals (Parylo et al., 2012). Parylo et al. (2012) wrote that principals felt the newer system was a lot clearer on what was expected of them. With the theme of dialogue, “principals conceptualized their evaluation process as a shared dialogue where they played an active role” (Parylo et al., 2012, p. 226). The principals in Parylo et al.’s study believed evaluation was dictated to the principals being evaluated in the past and now is more of a dialogue. The “trust and respect” theme had to do with the relationship the principals had with their evaluators. The principals trusted and respected the feedback received by their evaluators and the tools created and used by the district office to evaluate the principal. Because of the trust and respect given to the principal evaluators, feedback from the evaluators was also valued by the principals in Parylo et al.’s study. “Support” was the next theme presented by Parylo et al.. This support that the principals perceived had to do with the support that the evaluation system was supposed to provide to help principals improve their performance. Although the principals thought most of the changes to their evaluation systems were positive, the last theme mentioned by Parylo et al. was how the new system provided tensions. These tensions had to do with the increasing dependence on accountability and testing information, structure specific problems with the new system, and lack of professional development. Despite these perceived tensions, however, the
overall impression of the new evaluation system for these principals was positive (Parylo et al., 2012).

These four studies researched principals’ perceptions of their evaluations. Two of the studies (Johnson, 2005; Rouse, 2009) researched principals’ perceptions of the feedback they were receiving on their evaluations. The third (Muench, 2014) studied principals’ perceptions of a specific evaluation system. In the final study Parylo et al. (2012) gave a qualitative perspective to principals’ perceptions of their evaluation. Principals’ perceptions of their own evaluation system are important to help the system continuously improve (State Council for Educator Effectiveness, 2011).

**Literature review on principal evaluation.** Davis et al. (2011) reviewed 68 peer-reviewed and not peer-reviewed studies focused on principal evaluations. As a result of this literature review, Davis et al. (2011) came up with four key points that are best practices for principal evaluation:

- Rigorous empirical evidence regarding best practices in principal evaluation is extremely thin. As a result, it is difficult to assert the effects of evaluation on important school outcomes. Likewise, it is difficult to generalize effective principal evaluation practices found in one school district to all school districts.

- The quality of the conduct of principal evaluation may be more important than its content. Strong, trusting, and collaborative relationships between principals and their district office evaluators is especially critical to the success of the evaluation process.
• Establishing a balance between the formative and summative functions of evaluation appears to result in greater principal buy-in and motivation regarding the evaluation process.

• Principal evaluation systems appear to be most effective when they are based upon clear standards and expectations of performance and aligned with the key goals and needs of principals, schools, and districts. (p. 35)

The conclusions of Davis et al. (2011) connected well with the findings of this dissertation.

**Study specific to post Race to the Top state evaluation systems.** Fuller, Hollingworth, and Liu (2015) published a study of all 50 states’ plus Washington D.C.’s principal evaluation systems. They created this study considering recent federal legislation that “has created strong incentives for states to adopt principal evaluation systems, many of which include new measures of principal effectiveness” (Fuller et al., 2015, p. 164). They stated that with these measures came an accountability system that is heavily reliant on student, staff, and community outcomes. From their research came five themes. These five themes were:

• The purpose of these systems is to guide professional development to improve student outcomes,

• Sixty-eight percent of the states were using the results to guide high-stakes decisions,

• A majority of the states were collecting multiple forms of student outcomes as a part of the evaluation (not just high stakes tests),
Less than a majority of the states are using data showing a principal’s
direct influence (teacher, student and parent perception data),

There is little effort being put into evaluating the evaluation systems and
trying to improve them (twenty-three percent of all states)\textsuperscript{12} (Fuller et al.,
2015, p. 186).

From their research, Fuller et al. brought up many concerns about the data that was being
used to for principal evaluations and how the data was being used. Fuller et al. (2015)
concluded:

There are currently no strategies to estimate principal effectiveness that accurately
capture the independent effect of principals on student test scores; thus, these
current strategies send inaccurate signals to both principals and those who make
employment decisions about principals. (p. 466)

This conclusion and their concerns will be brought to light in the following section.

\textbf{Challenges Within Principals’ Evaluations}

The following paragraphs are a look into the challenges found by researchers to
principal evaluation systems. These challenges range from the amount and type of
feedback provided, to the inconsistent practice of systems, to the use of data in high
stakes evaluation systems.

Stronge (2013) and Ginsberg and Berry (1990) wrote of multiple challenges
within evaluation systems. Stronge (2013) concluded “many of the current principal
evaluation systems do not allow for shades of gray—principals are either rated as

\textsuperscript{12} Colorado Department of Education has a report on the validity of ratings which would
partially address the fifth theme.
http://www.cde.state.co.us/educatoreffectiveness/principal_validitystudy
satisfactory or unsatisfactory” (p. 6). Some systems were not aligned with professional standards, “which can produce role conflict and subsequent role strain as principals have trouble knowing what they should focus their attention on” (Stronge, 2013, p. 7). Poorly written parts of an evaluation can also be a problem (Stronge, 2013; Ginsberg & Berry, 1990). “Problems are seen with evaluation systems that have poorly stated criteria and standards of performance that leave questions in principals' minds about expectations” (Ginsberg & Berry, 1990, p. 220). As was mentioned previously, these same questions could occur in principals’ minds if their evaluators provide disconnected and/or poor feedback. One reason behind poor feedback may be that their evaluators have not been appropriately trained on the evaluation tool (Ginsberg & Berry, 1990). Per Stronge (2013), many systems had an absence of meaningful and timely feedback and are not coupled with consequences.

Researchers were not the only people who have concerns about evaluation systems, principals themselves also have concerns. Reeves (2004) detailed many opportunities for improvement in evaluation systems based on the results of a national survey of evaluation instruments to detail the challenges of evaluations. According to Reeves (2004), the following were the significant findings of the study:

- More than 18% [of the leaders] never received an evaluation in their current position.
- Eighty-two percent of the leaders who were evaluated said their evaluation was “inconsistent, ambiguous, and counterproductive”.
- Only 47 percent said their most recent evaluation related to student achievement.
• Only 54 percent believed their evaluation was based on clear standards.
• Forty-seven percent believed the evaluation was “sufficiently specific to help them improve their performance.”
• “The higher the level of leadership responsibility, the lower the satisfaction with leadership evaluation instruments.” (Reeves, 2004, p. 53)

For Reeves, these statistics pointed to the need for a better way to consistently evaluate educational leaders. A better evaluation instrument and process would be a process that gives better and more constructive feedback. “A better model would provide specific, accurate, and timely feedback. Rather than an event that occurs once a year, evaluation should consist of frequent feedback and provide multiple opportunities for continuous improvement” (Reeves, 2004, p. 57).

Fuller and Hollingworth (2014) researched principal effectiveness as it is judged using student test scores. The authors were not in support of student data being used for principal effectiveness. Fuller and Hollingworth argued that these tests were not originated to assess principal and teacher effectiveness and are not effective doing so. “Based on the research in this area, we can unequivocally conclude that even the most sophisticated and thoughtful efforts to estimate principal effectiveness are flawed and produce inaccurate results” (Fuller & Hollingworth, 2014, p. 491). Fuller and Hollingworth (2014) concluded that those efforts that best connect principal effectiveness to student data would be longitudinal where another principal’s impact could be less likely:

In sum, many – if not most – principals could not be included in the highest quality efforts to estimate effectiveness, the estimates themselves simply do not
accurately reflect the independent contributions principals make to student changes in test scores, and most states have adopted the most simplistic of efforts to estimate principal effectiveness. Thus, without question, using student test scores to estimate principal effectiveness is simply building a bridge too far. (p. 492)

The conclusions of Fuller and Hollingworth (2014) would contradict the guidelines of the CMES since 50 percent of the evaluation is based on student assessment data.

Fuller et al. (2015) expanded the questioning of the lack of legitimate data collection and analysis from not just student achievement and growth data, but to all statistics that are being used for principal evaluation. They stated that the data might be good for feedback to help with principal reflection. However, “there is little evidence that any approaches lead to conclusions about individual principals that are accurate, valid, and reliable enough to make high-stakes decisions” (Fuller et al., 2015, p. 187).

The first concern, had to do with using student testing data for evaluations, which was also mentioned by Fuller and Hollingworth (2014). Fuller et al. (2015) also had concerns with the use of teacher, student, and parent perception data for the evaluation of principals because of a lack of evidence proving it is accurate. Their next concerns address the use of teacher retention and quality as summative evaluation data. Fuller et al. (2015) concluded “there may not be any statistical method that can isolate the impact of a principal on the characteristics of teachers hired or retention of teachers” (p. 188).

Fuller et al. (2015) ultimately were questioning whether any of the data mentioned could be isolated for principal impact well enough to produce legitimate data pertaining only to the principal.
There have been challenges in evaluations systems that need to be corrected to benefit principals and their performance (Stronge, 2013; Ginsberg & Berry, 1990; Reeves, 2004). With some of the challenges coming from researchers who question the use of any of the data being used by current evaluation systems to provide legitimate information for high stakes decisions (Fuller & Hollingworth, 2014; Fuller et al., 2015). According to some researchers, challenges in evaluation systems point to the need for systems where frequent and timely feedback is provided to administrators to help them improve their performance (Stronge, 2013). Effective feedback will help principals improve (State Council for Educator Effectiveness, 2011).

A report of principals’ and principal evaluators’ perceptions of the CMES was conducted by the Colorado Department of Education after the state piloted CMES within certain school districts in Colorado (Colorado Department of Education, n.d.b). This report was found on CDE’s website and was presented as a report of findings from a survey they sent to pilot principal’s and their evaluators. CDE created a survey that was given to principals and principal evaluators (who were a part of a pilot – convenience sampling) asking for feedback about the CMES (Colorado Department of Education, n.d.b). The first survey (fall 2011) gauged satisfaction with the pilot school districts’ pre-CMES evaluation systems. The second survey (spring 2012) gauged satisfaction with the pilot CMES system. The third survey (spring 2013) gauged the satisfaction with the CMES pilot system, again.\textsuperscript{13} According to the answers given by survey respondents, principals felt more positive about the CMES system when compared to their previous evaluation systems (Colorado Department of Education, n.d.b). The smallest difference

\textsuperscript{13} Survey results can be found in Appendix C.
in percentage between old system and first pilot was +26 percent and the largest difference in positive responses was +55 percent. This positive feeling again increased from first year in the pilot to second year in the pilot (2012 spring – 2013 spring) with the smallest percentage growth for one theme being six percent and the largest growth in theme being twenty-four percent (Colorado Department of Education, n.d.b). According to C.D.E. the CMES system was more satisfactory to principals and principal evaluators than their previous systems, and it improved with changes made by C.D.E. to the system and with evaluators’ continued experience in the system (Colorado Department of Education, n.d.b).

**Role(s) of the Evaluators of Principals**

In an article by NAESP and NASSP (2013) it was stated, “The quality of how principal evaluations are conducted may be even more important than the content of the evaluations” (p. 32). An important part of the implementation is the evaluator/ supervisor of principals. Eight standards were released by the Council of Chief State School Officers (CCSSO) as guidelines for supervisors of principals to define their jobs (CCSSO, 2015). These standards are optional for school districts to use, but will help to provide clarity to the jobs of those who supervise principals (Superville, 2016).

In particular, the standards emphasize the supervisors’ role in helping the principals they oversee improve as instructional leaders; in serving as liaison between schools and the central office; and the supervisor’s own responsibility to grow as a leader (Superville, 2016, p. 4).

According to the Council of Chief State School Officers (CCSSO) (2015) the eight standards cover the following professional practices for principal supervisors:
• Standard One – Principal Supervisors dedicate their time helping principals grow as instructional leaders.

• Standard Two – Principal Supervisors coach and support individual principals and engage in effective professional learning strategies to help principals grow as instructional leaders.

• Standard Three – Principal Supervisors use evidence of principals’ effectiveness to determine necessary improvements in principals’ practice to foster a positive educational environment that supports the diverse culture and learning needs of students.

• Standard Four – Principal Supervisors engage principals in the formal district evaluation process in ways that help them grow as instructional leaders.

• Standard Five – Principal Supervisors advocate for and inform the coherence of organizational vision, policies and strategies to support schools and student learning.

• Standard Six – Principal Supervisors assist the district in ensuring the community of schools with which they engage are culturally/socially responsive and have equitable access to resources necessary for the success of each student.

• Standard Seven – Principal Supervisors engage in their own development and continuous improvement to help principals grow as instructional leaders.
• Standard Eight – Principal Supervisors lead strategic change that continuously elevates the performance of schools and sustains high-quality educational programs and opportunities across the district. (CCSSO, 2015, pp. 14-22)

Superville (2016) quoted Jody Spiro of the Wallace Foundation on the importance of the principal’s supervisor position and therefore the standards: “If you get that position right it’s beginning to become clear that it has quite a big effect, because the schools can’t do business as usual and the central office can’t do business as usual” (as cited in Superville, 2016, p. 5). This also ties to the evaluation process. In a review of literature on principal evaluation, Davis et al. (2011) found “strong, trusting, and collaborative relationships between principals and their district office evaluators is especially critical to the success of the evaluation process” (p. 34). This stress on the importance of the supervisor of the principal also highlights the importance of the feedback given to principals by their supervisors. With clear indicators (through standards) as to the expectations of roles that supervisors must play (Superville, 2016) supervisors will have an opportunity to provide clear and effective feedback to those principals they supervise.

**Research on Feedback Intervention Theory and Regulatory Focus Theory**

FIT and RFT are two theories highlighted in the first chapter to provide a theoretical framework for feedback. Both theories and the studies that provide the foundation for both of these theories have been sited quite often in research. According to Google Scholar, the original article that Kluger and DeNisi (1996) wrote as a foundation for FIT has 4300 citations currently (Google Scholar, n.d.). The original article by Brockner and Higgins (2001) on RFT has 564 citations (Google Scholar, n.d.).
During this section of the literature review the researcher will focus on a few articles that specifically research both theories.

Research on Feedback Intervention Theory

Most of the studies reviewed referenced FIT and parts of FIT to provide some context on feedback. Some articles study specific parts of FIT, however. The following three articles were chosen for various reasons. Hergovich et al. (2013) were chosen because of their specific focus on FIT and the impact of feedback on goal setting and task performance. Martin and Mottet (2011) tied FIT into their research on feedback that students are receiving in a K-12 setting. Finally, Khachatryan (2015) tied FIT to supervisor feedback between teacher supervisor’s and teachers. All three of these studies combined members of authority in education (professors, teacher supervisors, teachers) to subordinates (college students, teachers, high school students) who needed to identify some growth that needed to be done to improve performance and an observation of how feedback was given using FIT as the model.

Hergovich et al. (2013) created a research study in which college students were given tasks and feedback. The parts of FIT that Hergovich et al. (2013) was testing were positive and negative feedback and the feedback’s impact on task performance. The study found that people mostly raised their performance when the standard was met and maintained their performance when the task was not met and negative feedback was received (Hergovich et al., 2013). Unlike what is hypothesized by Kluger and DeNisi (1996) the effort was not increased for those people who received negative feedback verses those who were receiving feedback that caused them to increase the standard (Hergovich et al., 2013). In fact, those who received constant negative feedback, either
maintained their effort, lowered their performance effort, and/or demonstrated avoidance behaviors (Hergovich et al., 2013). In this study, negative verses positive feedback seemed to make an impact on effort and performance.

Martin and Mottet (2011) specifically tested FIT’s assumptions around feedback’s effectiveness, when given indirectly, connected to goals and was relational in nature, verses feedback that had more to do with the task. Their test subjects were Hispanic students, and the relational feedback was related to the feelings of those receiving the feedback and the task feedback had to do with a specific writing task. This study did not have the students receiving feedback, they were given written scenarios. Martin and Mottet did not find a significant impact when comparing feedback that related to the student’s feelings verses feedback related to the task. Martin and Mottet did find that rapport the teacher builds or has built with the students does have significant impact. This research did not promote or counter FIT, and in their findings, the authors made suggestions to better connect their work to FIT (Martin & Mottet, 2011).

Khachatryan (2015) used FIT to examine the relationship between feedback given by teacher supervisors to teachers based on classroom observations and the teachers’ perceptions of feedback. The attention of this study was directed to the three potential ways in which attention can be directed when feedback is given: focus on the self, focus on the focal task, and focus on the task details. Khachatryan found that 13 percent of the feedback was specifically focused on the self, and that the vast majority of the feedback focused on the focal task or focal processes. When teachers’ perceptions were considered, three themes arose from the data: most of the feedback felt validating, many feedback comments caused teachers to reflect on their practice, and some comments were
taken with skepticism (Khachatryan, 2015, p. 177). These themes facilitated teachers reflecting on their practice. For teachers “feedback that focuses attention on the details of instructional moves—the “task details” in FIT terms—prompts teachers’ learning processes and plans for changes in teaching” (Khachatryan, 2015, p. 183). This research, though done with a small sample size, would reaffirm the need for effective feedback to be on the focal task and task processes.

These three studies, that used parts of FIT, all had a reason for being a part of the literature review. The first was that all three put those receiving feedback into a scenario where they were the subordinate. Hergovich et al. (2013) used a controlled environment in order to compare positive to negative feedback and the implications on effort and performance. Hergovich et al. (2013) did not consider, when looking at the impact, the human side of delivering this impact. Martin and Mottet (2011) tested the assumption by FIT that feedback is more productive when given related to the task as opposed to related to the self, and found no difference between the two. This study did not put the students into an actual feedback scenario. Finally, Khachatryan (2015) was placed into this literature review because it gave a scenario where supervisors were giving feedback to subordinates, and the study used FIT as a basis for its analysis. Overall, these three studies were used to show how FIT was being researched.

**Research on Regulatory Focus Theory**

A scan of the research around Regulatory Focus Theory ties mostly into work environment and feedback within the work environment. The first study evaluated is a meta-analysis based on RFT in work settings (Apodaca et al., 2012). For Carlson et al.’s (2008) study the researcher looks into leadership and how subordinates react to different
styles of leadership using the RFT perspective. The third study on RFT was co-authored by one of the originators of FIT, Kluger, and provided at least a partial integration between the two theories (Dijk & Kluger, 2011). From these three studies comes an understanding of the research around RFT.

Apodaca et al. (2012) wanted to answer four questions in their meta-analysis of RFT. The first question asked about the ways that promotion verses prevention focus were measured (Apodaca et al., 2012). The result of this question was 14, which the researchers felt was too many measures and that a more consistent way to measure needed to be determined. The second question asked about the empirical relationship between promotion and prevention focus. “Assuming perfect reliability, promotion and prevention focus are at best weakly related and appear to be distinct and orthogonal factors” (Apodaca et al., 2012, p, 168). Apodaca et al. believed that this provided a good foundation for research around promotion and prevention focus. The third question asked about promotion and prevention focus and how they connect with work related variables. Apodaca et al. (2012) found that “promotion focus was related to positive individual differences and work outcomes, and that prevention focus was related to negative individual differences and work outcomes” (p. 168). The final question asked about whether RFT predicts job satisfaction. Apodaca et al. (2012) “found that regulatory focus does explain a significant proportion of unique variance in job satisfaction and task performance” (p. 169). The conclusion that the researchers came up with was that RFT was worth continuing to study.

Carlson et al. (2008) researched the mediating impact of leadership behavior on employees using the RFT perspective with two different leadership styles in mind:
Initiating Structure and Servant Leadership. The results of the study connected initiating structure leadership to deviant and in-role behaviors therefore confirming prevention focused behavior (Carlson et al., 2008). “By modeling initiating structure, leaders induce in employees a focus on security, obligations, and loss avoidance” (Carlson et al., 2008, p. 1229). Servant leadership was associated with helping and creative behaviors therefore connecting to promotion focused behavior. “By modeling servant leadership, leaders induce in employees focus on nurturance, aspirations, and gains” (Carlson et al., 2008, p. 1229). Carlson et al. added a new dynamic to RFT they call Work Regulatory Focus (WRF) which attempts to account for changes in regulatory focus as opposed to Higgin’s work which accounts for chronic regulatory focus. WFR differs from RFT because WFR captures “the psychological state of an employee at any point in time” (Carlson et al., 2008, p. 1229). This way of looking at regulatory focus provided a more responsive measure to employees’ mindset (Carlson et al., 2008).

In the following research, one of the authors was one of the originators of FIT (Kluger). Dijk and Kluger (2011) focused on task type and whether or not the task type made feedback feel positive or negative. Dijk and Kluger did a pre-test and two studies on these topics and came up with two key findings. First is “task type activates regulatory focus” (Dijk & Kluger, 2011, p. 1099). Those tasks which were presumed to be prevention focused were perceived to be that way by the sample used for the research. The same held true for promotion focused tasks. The second finding was “task type moderates feedback-sign effects on motivation and performance” (Dijk & Kluger, 2011, p. 1099). People performing tasks which would be considered prevention focused were more receptive to negative feedback than positive feedback and the opposite was true for
those performing promotion based tasks (Dijk & Kluger, 2011). This study combined an important component of both FIT and RFT.

These three studies provide a good perspective on research studying RFT. Apodaca et al.’s (2012) meta-analysis provided a general outline of how RFT was being researched, and the tools used to research RFT. Carlson et al. (2008) used RFT to examine promotion verses prevention focused behavior and used it to create a new way to gather data around regulatory focus, WFR. Finally, Dijk and Kluger (2011) studied RFT and a basic principle of FIT and found a relationship between motivation and the perceived regulatory focus of the task. The studies of FIT and RFT provide a strong foundation for the researcher of this paper to look into feedback and the perceptions of this feedback. This research, along with the rest of the research in this literature review provided a strong foundation of research for this study. The gap in the research is explained below.

**Gap in the Research**

Principals’ indirect effects on student achievement in schools are critical to the improved performance of their schools (Ahtaridou et al., 2009; Anderson, Leithwood, Seashore Louis, & Wahlstrom, 2004; Coelli & Green, 2012; Hallinger & Heck, 1996; Marzano et al., 2005). Many researchers have identified skills, responsibilities or roles that a principal must have or perform to improve school performance (CCSSO, 2008; Clifford et al., 2012; Fullan, 2002; Gold, 2003; Marzano et al., 2005; Matthews & Crow, 2010; Hallinger et al., 2013; Williams, 2008). Through the last century, expectations of skills, roles, and responsibilities of principals have shifted to focus on instructional leadership and is being enforced through evaluation systems (Ginsberg & Thompson,
Race to the Top explicitly intended to impact evaluation by making student achievement and outputs a primary focus of evaluations, and asking that applicants (States) make significant shifts in their educator evaluations to win money provided by the grant (McGuinn, 2011; US Department of Education, nd; US Department of Education, 2009).

Colorado was one of the states whose Governor and legislative branch adopted legislation to change the evaluation system for educators in Colorado (C.R.S. 22-9-105, 2013; Executive Order B 2010 01, 2010).

With the adoption of the Educator Effectiveness Act in Colorado in 2010, it was legislated that “the state will develop a high-quality, implementation-ready model evaluation system, with associated tools, available to any district that chooses to use the model system” (State Council for Educator Effectiveness, 2011, p. 9). With input from many educators, a model evaluation system was developed and piloted. Given the option of either adopting the state’s model evaluation system or developing a comparable system of their own, most districts in the state of Colorado adopted the state’s model evaluation system to evaluate all of its educators (including principals).\(^{14}\) This new system provides a researcher with many principals who are transitioning from multiple different systems into one uniform system. This system has a single rubric that defines the expectations, traits, and behaviors of principals (Colorado Department of Education, 2016-2017).

The CMES also provides the opportunity for a large group of evaluators to have a unified purpose. According to the assumptions of the Council “the purpose of the system

\(^{14}\) Out of the 192 school districts or BOCES in Colorado, 174 are using the State Model teacher evaluation system and 164 are using the State Model System for principals.
is to provide meaningful and credible feedback that improves performance” (State Council for Educator Effectiveness, 2011, p. 7). Since “providing meaningful and credible feedback to improve performance” is the purpose of the system, then it would be important to ask those who are receiving the feedback in this system about the meaningfulness and credibility of the feedback they are receiving. In this study, those receiving the feedback were school principals. It is also important to gather this information to look at the perceived priorities of feedback given by principal evaluators. Finally, with many districts using the rubric provided by the state of Colorado and the unified purpose of this evaluation being professional growth through substantive feedback, the strengths and opportunities for improvement of the system might inform those who are trying to make the system better. With continuous improvement of the CMES as a stated assumption, a study on feedback could help to improve the CMES.

Researchers who studied FIT and RFT provided a strong foundation for understanding of these theories. Researchers studying FIT looked mostly into feedback related to the self verses feedback related to the task and how people react to these different levels of feedback (Hergovich et al., 2013; Martin & Mottet, 2011; Khachatryan, 2015). Researchers studying RFT focused on promotion verses prevention focused behaviors by people who would be considered subordinates (Apodaca et al., 2012; Carlson et al., 2008; Dijk & Kluger, 2011). One of the sets of researchers looked specifically at constructs that are important in both RFT and FIT (Dijk & Kluger, 2011).

The researcher of this study studied the perceptions of the feedback received as a part of the CMES. Principals were specifically identified as those who were receiving
feedback for their own evaluations and their perceptions on whether or not the feedback altered their practice.

The research questions that guided this study are:

Q1 What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process?

Q2 What are principals’ perceptions of the CMES evaluation feedback?

Q3 According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

The findings of this study benefit evaluation systems for principals by providing a better understanding of principals’ perceptions of the feedback they are receiving on the CMES.
CHAPTER III

METHODODOLOGY

Including the introduction there are seven sections to Chapter III. The researcher will present both the overview and the goal of the research in the following section. The overview will be followed by the research questions. Following the research questions will be the research design. This research design will have the philosophical perspective of the researcher for this research project and the actual research design that will be used by the researcher to answer the research questions. After the research design is presented, a detailed step-by-step process of the research methodology will be outlined in order to specify the use of the instrument for data collection. Next, the researcher bias will be brought into light based on the researcher’s experience with the CMES system. The researcher makes certain assumptions, which presume the type of research and the research instrument will provide the most successful path for this study. At the conclusion of Chapter III, the researcher hopes the reader of this dissertation has a thorough understanding of the research completed by the author.

Overview

The authors of Race to the Top sought to, among other efforts, improve principal and teacher performance through the evaluation process. Lawmakers in the state of Colorado followed the direction of the federal government by creating the Executive Order B 2010-001 and the Educator Effectiveness Act. These two pieces of legislation were to change the evaluation process and the instrument used for the evaluation process.
in Colorado. The lawmakers tasked the State Council for Educator Effectiveness (The Council) with creating the framework for the new evaluation system. The Council then determined the primary purpose of the evaluation system is “to provide meaningful and credible feedback that improves performance” of teachers and principals (State Council for Educator Effectiveness, 2011, p. 7). If an evaluation system’s purpose is to provide feedback to improve the performance of those being evaluated, then those being evaluated must be able to perceive whether or not they are improving as a result of the feedback.

Out of the 192 school districts or BOCES in the state of Colorado, 174 are using the Colorado Model Evaluation System (CMES) for teachers (Colorado Department of Education, 2016b). In addition, 164 of these school districts or BOCES are using the CMES for principals.\(^\text{15}\) There are currently a large number of school districts using the Colorado Model Evaluation System.\(^\text{16}\) This number provided an opportunity to get a large sample of principals to answer a survey instrument.

If the purpose of the Colorado Model Evaluation System is to provide feedback to educators so that they might improve their performance, then what are the perceptions of the feedback that is being given to principals for their evaluation? Given the fact that the Colorado Model Evaluation System for principals was formally used for the first time in the 2014-2015 school year, there is a need to document principal perceptions on

\(^{15}\) The districts who are choosing not to follow the normal framework provided by the Colorado Department of Education, would be obligated to create their own. “Any district that chooses to develop its own system may do so, provided that mandatory elements required for educator evaluation systems are included and state technical guidelines are met” (State Council for Educator Effectiveness, 2011, p. 9).

\(^{16}\) The list of districts or BOCES using this instrument can be found at the following web address: http://www.cde.state.co.us/educatoreffectiveness/assurancesreport_2016
how the tool (and feedback using the tool) is changing practice to better gauge its overall effectiveness.

**Research Question**

Although the people giving the feedback to principals and making the final evaluation rating to judge performance are evaluators of principals, the principals are those whose performance must be improved. Documenting principals’ perspectives of this evaluative feedback helps to understand the perceptions of the feedback they are receiving. With a large number of principals in the state of Colorado being evaluated using the same tool, finding common answers to their perceptions of the Colorado Model Evaluation System and the feedback they are receiving provided an opportunity to get an understanding of these principals’ perspectives on the following research questions:

- **Q1** What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process?
- **Q2** What are principals’ perceptions of the CMES evaluation feedback?
- **Q3** According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

The perspective taken by the researcher and the research design used for this study are explained below.

**Research Design**

The researcher aimed to get a large number of principals in the state of Colorado to give their perspectives on feedback using the CMES. With the resulting sample of principals, the trends of these principals’ perceptions were then determined using this information. Since the researcher wanted perceptions of feedback using the CMES and opportunities for principals of the sample to expand on their answers, mixed methods
research design was chosen. The researcher had the assumption that “the uses of both quantitative and qualitative methods, in combination, provide a better understanding of the research problem than either method by itself” (Creswell, 2012, p. 535).

Pragmatism – A Philosophy for Mixed Research

Hibberts and Johnson (2012) stated that pragmatism is a research philosophy that is often used in mixed research. Pragmatism means “you should mix research components in ways that you believe will work for your research problem, question, and circumstances” (Hibberts & Johnson 2012, p. 124). Hibberts and Johnson further explained pragmatism as an attempt to find a middle ground between philosophical dualisms. It “views knowledge as both a human construction and a product of the reality of the world we live in” (p. 125). Finally, Hibberts and Johnson (2012) stated that those who believe in the philosophy of pragmatism, “view organisms as continually adapting to their worlds. The present is always an opportunity to improve understandings in a way that fits and works in our physical, social and psychological environments” (p. 125). These connections all are important to the success of mixed research (Hibberts & Johnson, 2012).

In this particular study the decision to use mixed research by the researcher worked well for the research problem, question, and circumstances. Principals’ perceptions of the CMES take into account both how principals construct their knowledge of the system and the realities of the system that has been constructed to ensure they are evaluated. Principals must be adaptable to succeed. The CMES is supposed to be a system that is adaptable and, ultimately, improving because it is supposed to be continuously improving. Finally, the survey instrument used for this
research study followed the pragmatic philosophy because the construction of this instrument allowed principals to answer questions with Likert answers followed by questions that allowed the principals to expand on the Likert answers they gave. This allowed the researcher the opportunity to build an instrument “that works for the research problem(s), question(s) and circumstances” (Hibberts & Johnson, 2012, p. 124).

**Methodology – Mixed Methods Using Survey Design**

From the foundation provided by this researcher’s philosophy, the next step in the research process will be the methodology (Hibberts & Johnson, 2012). The methodology is “the strategy, plan of action, process or design lying behind the choice and use of particular methods and linking the choice and use of methods to desired outcomes” (Crotty, 1998, p. 3). The researcher used a mixed methods approach sending and online survey questionnaire through *Qualtrics*. According to Creswell (2012) the type of design used is a convergent parallel design where both qualitative and quantitative research designs are converged in order to answer the research question. Having principals complete a mixed methods survey allowed the researcher to capitalize on strengths of both the qualitative research and the quantitative research. Most notable strengths for mixed methods research is: “convergence and corroboration of findings can enhance evidence of a particular claim” (Hibberts & Johnson, 2012, p. 126).

This design allowed the researcher the opportunity to do an environmental scan of Colorado principals’ perceptions of the CMES system and the feedback they received using the CMES for Principals and gave the principals the opportunity to expand on their thinking. The principals’ perspectives on if their practice was being altered by feedback using the CMES was analyzed from the results. The gap in the research is determining if
the third assumption according to the Council is being fulfilled according to the perceptions of principals and how these perceptions might help improve CMES. Because of the gap in research tying feedback received by principals to the evaluation tool that is supposed to have a formative and summative purpose, an environmental scan was done.

Initially, the survey was piloted using assistant principals. This was then followed by a final survey which was edited and given to principals. The survey questionnaire had 24 questions total. The first 13 questions were demographics questions and the consent to use the data collected. The final 11 questions included three questions that had Likert Scale responses ranging from: extremely positive to extremely negative with a neutral option in the middle. These gave the participants five options for questions with Likert Scale answers. The remaining questions were qualitative in nature and boxes were provided to answer the questions as thoroughly as the principals wished to answer. All of the qualitative questions were asked as follow up questions to the closed questions with Likert responses.\(^\text{17}\)

**Population.** The population is a “group of individuals who have the same characteristic” (Creswell, 2012, p. 142). This target population was all public-school principals who are being evaluated using the Colorado Model Evaluation System (CMES) for principals.\(^\text{18}\) *The Educator Effectiveness Act* allowed for public school districts or Board(s) of Cooperative Educational Services (BOCES) to build their own instrument with the CMES for principals as the model or choose to use CMES as it was and is continuously being developed by the Colorado Department of Education (CDE).

\(^{17}\) Final survey given to principals can be found in Appendix D.  
\(^{18}\) The population for the final research paper will include K-12 principals. K-12 assistant principals will be included in the pilot population only.
The target population for this study was the public-school principals who are being evaluated using the CMES for principals. The researcher of this study invited the entire population to take the online survey questionnaire.

The email addresses for the potential participants were purchased from a website that was supposed to have all principal contacts in the state of Colorado. The name of this website is www.emaillistus.com. This company was used because the researcher noticed that many district websites did not have email addresses of their administrators on their website, a fellow doctoral candidate used this list before, and purchasing the list greatly reduced the amount of time used trying to locate email addresses. All principals on this list were asked to take the survey. The assumption made by the researcher was that these principals will know whether or not they are being evaluated through the CMES or took the time to look at the link provided by the researcher if they did not know if they were being evaluated using CMES.

Survey research design. Using survey research design best answered the research question. Through this research study the perceptions of the entire population of public school principals who are being evaluated using the CMES for principals was solicited. To best find the population’s perceptions, Creswell (2012) stated “the attitudes, behaviors, opinions, or characteristics of an entire population are determined through a survey of either the entire population or a representative sample of it” (p. 376). The following information will first give the issues in which it is best to use survey research (Muijs, 2012). After this, the identification and description of the survey will be outlined. This information will start with the type of design (cross-sectional), then identify the
variables that will be researched, followed by the detailed description of the creation of the instrument. (Creswell, 2012)

**When to use survey design.** Muijs (2012) gave four types of issues in which survey research is the best research method to be used. These four issues are: “when the issue is about quantity; when the issue is about opinions and attitudes; when the issue is about relationships between variables; and when the issues are about perceived behaviours” (p. 141). In asking the main research question, the issues studied included quantity, opinions and attitudes, and perceived behaviors. The number of principals who are being evaluated using the same evaluation system (CMES) is significant. Most of the districts in Colorado are using this system. This quantity of principals makes a survey a great research instrument. The research question asked the perceptions of the principals of the feedback they were receiving using the CMES for principals. Their perceptions were their opinions and attitudes of the feedback. The perception of feedback received from an evaluator is a perceived behavior of the evaluator or the system of evaluation by the principals.

**Cross-sectional survey design.** The two choices of survey designs presented by Creswell (2012) are longitudinal and cross-sectional design. This research study was approached using the cross-sectional design. This design grabbed data at one specific point in time. By being able to identify one specific point in time using the cross-sectional design, a researcher “can examine current attitudes, beliefs, opinions, or practices” (Creswell, 2012, p. 379). Within the cross-sectional design there are three different approaches: attitudes and beliefs, community needs, or program evaluation.
Cross-sectional design was used for this research project to get principals’ perceptions of the feedback at a particular point in time. The reason why feedback was prioritized in this study is because one of the main purposes of CMES is to use feedback to improve the performance of those who are being evaluated using the system (State Council for Educator Effectiveness, 2011). Since the principals are supposed to be the ones within the system who are to improve because of the feedback they are receiving, then getting their attitudes and beliefs on the feedback would be an important understanding to have. So, cross-sectional design surveying the attitudes and beliefs of public school principals being evaluated using CMES was appropriate to use to answer the research questions.

Creating the survey. According to Creswell (2012), questionnaires “are forms used in survey design that participants in a study complete and return to the researcher. Participants mark answers to questions and supply basic, personal, or demographic information about themselves” (p. 626). As was stated earlier, the researcher created the survey instrument using both open and closed-ended questions. This survey was piloted with assistant principals who were being evaluated using the CMES prior to sending out the final survey. The initial questions were demographics questions. These questions were then followed up by a mix of open and closed ended questions that asked for principals’ perspectives of the three research questions. Each page of the survey started with a closed question requiring responses from a range of Likert answers. These closed questions were then followed by open ended questions from the same topic as the closed questions.
**Type of survey questions.** The researcher for this study wished to find personal information about, attitudes of, and behaviors of the principals who are being evaluated using the CMES. These three types of questions are important in survey questionnaire construction according to Creswell (2012). The personal information assesses the characteristics of individuals in the sample through the answering of demographics questions. The attitudinal questions are measures by which researchers measure the feelings toward educational topics by individuals. The final type of questions that were asked in this survey questionnaire was behavior questions (Creswell, 2012). These questions asked about the behavior of the evaluators of principals. Using these three types of questions, a survey questionnaire was created.

Much thought was put into the demographics used to break down the trends in the information collected for the other questions. Demographics were used to compare the sample to the data collected by CDE, and/or to see if there were any discrepancies in answers given for the three closed questions with Likert Scale answers. The demographics collected are listed below, along with reasons behind using these demographics:

1. Gender of the principal (Male or Female). This information is collected by CDE.
2. Years as a principal. (Blank space left for the principal to fill in the number) Principals years in a position might be a factor in their perception of their experience with evaluation.
3. Years in current position. (Blank space left for the principal to fill in the number) This potentially ties into Coelli and Green’s (2012) research.
4. Type of district (Pre-categorized - Urban, rural, suburban). CDE collects this data, but uses different categorization.

5. Location of district (Pre-categorized – Northern Mountains, Central Mountains, Southern Mountains, Western Slope, Eastern Plains, Front Range).

6. Free and Reduced Lunch Rate of school (Blank space left for the principal to fill in the number) Principals’ experiences with the evaluation tool might differ according to FARM rate of school.

7. Size of school – (Built this question with answers in 200 student increments) Participant data was compared to CDE data for this demographic.

8. Principal Evaluator type – (Pre-arranged list – Superintendent, Assistant Superintendent, Director of Secondary or Elementary, Other). These are the people who will evaluate principals. Roles of the evaluators outside of their evaluation duties can impact the amount of time given to the evaluation process and implementation of the system of evaluation.

9. Gender of the evaluator (Male or Female). The experience of the evaluator and the principal might be different according to either or both of their genders.

10. Type of school – (Pre-arranged list – elementary, middle, high, junior high, K-8, K-12). Principals in these different types of schools may view the system differently because of how the structure of their school has
impacted the system as a whole. CDE also had comparable data to the sample data collected.

Survey data was disaggregated based on principals’ demographic data.

Attitudinal questions were the majority of the questions on the survey questionnaire. The reason behind the use of attitudinal questions was because principals’ perceptions of the feedback given using both the evaluation tool as whole, and individual parts of the evaluation process and tools were being researched. These perceptions were ultimately asking for the attitudes of the principals towards the CMES system, the feedback they are receiving, and whether the feedback on the systems altered the principals’ practice, therefore attitudinal questions were asked.

Each closed-response question had a range of responses that followed Likert’s scale. Likert’s scale is “a scale used to represent people’s attitude to a topic” (Stevenson, 2015, para. 1). It was created by and named after its creator Rensis Likert, an American psychologist (Stevenson, 2015). When investigating individuals’ attitudes towards certain topics or things, “an individual may have a positive or negative predisposition toward an object and may hold this predisposition with more or less strength” (Alwitt & Berger, 1993, para. 6). With this in mind, Likert created a way to gauge whether or not people agree or disagree with something and also how strongly they believe their convictions:

When a Likert scale is used to measure attitude, its usual or standard format consists of a series of statements to which a respondent is to indicate a degree of agreement or disagreement using the following options: strongly agree, agree, neither agree nor disagree, disagree, strongly disagree. (Albaum, 1997, p. 332)
Because of the dual purpose of finding whether or not someone agrees with something and the strength of that persons’ convictions, the majority of the questions for this survey questionnaire were asked with Likert’s scale type answers.

There was one behavioral question asked on the survey. The reason behind the question was to find out the behavior of the evaluators of the principals, not of the principals themselves. The behavior of the evaluators that were being asked about was the frequency of evaluative feedback given to the principals.

From the personal, attitudinal, and behavioral questions, the researcher hoped to receive information that will help to find the answer to the research questions. From the foundation of these three types of questions, the format of the survey instrument needed to be developed. This format is described below.

**Instrument creation.** When creating this instrument, thought was put into what would be the most effective format to use in order to get a higher response rate from the principals offered the survey. The survey was a questionnaire asking questions that were both qualitative and quantitative using a web-based program called Qualtrics. The researcher strived to create a survey that maintained a balance between being thorough enough to provide rich and meaningful information and brief enough to allow for principals to complete the survey without taking too much time.

According to Creswell (2012) the advantages of using a web-based design are: speed of the responses, more people have access to computers, and it allows for a survey of the entire population. The researcher also believes that online instruments are best for principals because of the continued focus in schools on 21st Century skills and the tools
of the 21st Century. With all of these positive connections, a web-based survey was determined to be the best way to gather the needed information.

According to Creswell (2012) the disadvantages to using a web-based survey tool are as follows: According to the filters of each email system, many emails might be sent to spam. The survey would also be dependent on the reliability of the internet and internet browsers of those receiving the email. Coupled with the potential internet issues might be the potential for issues with the computers or devices that the principals use themselves. Finally, though technology is promoted in many school systems, making a survey questionnaire that can only be taken online creates a bias towards those who are computer savvy or those who have computers.

**Conclusion to building the survey instrument.** The way in which this survey questionnaire was distributed to all public school Colorado principals being evaluated using the CMES was through a web-based questionnaire using the Qualtrics system. This survey questionnaire was created using the guidelines set by the Institutional Review Board through the University of Northern Colorado. Permission of each principal was given or denied when the principal taking the survey answered the first question on the survey questionnaire. Principals taking the survey were also allowed to opt out of the survey at any time while they were taking the survey questionnaire. A principal was then taken through ten demographics questions. This was followed by the one behavioral question. The last questions on the survey questionnaire were attitudinal questions, which had a mixture of closed response (quantitative) and open-ended (qualitative) questions.\textsuperscript{19} After creating the survey, it was important to make sure the survey was valid.

\textsuperscript{19} The survey is Appendix D.
and reliable. For the quantitative research of this study, the validity and reliability of this survey was attempted using the methods highlighted in the following paragraphs. After the analysis of how validity and reliability is explained, the paragraphs following will explain how this study looked into reliability, triangulation, credibility, and transferability for the qualitative research.

**Reliability, triangulation, credibility, and transferability.** Determining the reliability, triangulation, credibility, and transferability of a research process are important components to the research project (Bush, 2012; Creswell, 2012; Merriam, 2009; Roberts, Priest, & Traynor, 2006). Following this paragraph, the setup and determinations made for the final survey because of the pilot study are explained. This is followed by explanations of how different reliabilities were determined. This is followed by how triangulation was used to assess the authenticity of the research. The final paragraphs contain explanations of how transferability and credibility were determined for the qualitative data.

**The pilot survey.** The survey was piloted using assistant principals being evaluated using CMES and then adjusted according to the results and conclusions made by the researcher. A pilot is “a procedure in which a researcher makes changes in an instrument based on feedback from a small number of individuals who complete and evaluate the instrument” (Creswell, 2012, p. 625). The method and results of the pilot are detailed in the paragraphs below.

According to Creswell (2012), the sample for the pilot study should not be drawn from the population who will be given the final survey. With this in mind, an alternative population was used for the pilot study. For the CMES, assistant principals and
principals go through the same process when being evaluated (Colorado Department of Education, 2016a). This includes using the same rubric (Colorado Department of Education, 2016a). Since the assistant principals and principals are evaluated using the same system, assistant principals being evaluated using the CMES were used for the pilot sample.

The pilot survey was sent to a list of assistant principals known by the researcher from both personal and professional experiences and to random assistant principals who could be found through various school district websites.20 Twenty-seven assistant principals were sent the survey. Sixteen of these assistant principals completed the survey. At the end of the survey there were three follow up questions.21

As was suggested by Creswell (2012), the researcher chose Cronbach’s Alpha to determine the internal reliability of the 16 responses for the two research questions that had Likert responses. The Cronbach’s Alpha was .775 for these two questions. This then meant that the two questions with Likert responses had a high internal consistency.22

When running a pilot study, Muijs (2012) had some things to look for that might suggest questions that need to be adjusted or excluded from the final survey. These suggestions included: look for anomalous answer patterns that suggest the questions

---

20 From this experience, the researcher noticed that many school districts did not have the email addresses of their administrators on their websites.
21 These three questions were: How many minutes did it take for you to complete the survey? Were there any questions in particular that were confusing or difficult to answer? (Please explain) 3. Please provide any suggestions on how this survey might be improved.
22 Although the standards for what makes a “good” α coefficient are entirely arbitrary and depend on the theoretical knowledge of the scale in question, many methodologists recommend a minimum α coefficient between 0.65 and 0.8 (or higher in many cases); α coefficients that are less than 0.5 are usually unacceptable (Goforth, 2015).
might not have been understood; Look for questions that might not be answered by a large number of respondents; asking the respondents to complete a second survey asking about the answers; (and/or) interview the respondents (Muijs, 2012). Creswell (2012) suggested that people in the pilot sample be allowed to write comments explaining any concerns the respondents might have had when responding to the survey.

Muijs’ (2012) suggestions were taken into consideration when the researcher evaluated the answers to the pilot survey. Anomalous patterns of answers came in three forms: responses were recorded that related to CMES for teachers, an extremely low number of answers were given to a final question, and the answers to the final research question included short and non-descriptive answers. First, the researcher noticed that a few of the assistant principals were answering the questions as if the questions were asked about the CMES for teachers. Language was changed in the initial description and the researcher also made sure to stress in reminder emails sent to the potential participants for the final survey that the survey questionnaire was about their own evaluation. In the pilot survey, a final question was included that asked if there was anything else the participants wanted to add. This question was excluded from the final survey, because only 6 people responded to the question with four of the responses being a version of “no.” In the pilot survey, the final research question was asked in a qualitative format. Thirteen of the 16 pilot participants responded to this question. Many of the answers to this question were short with the largest number of answers stating that the feedback has not impacted their practice. For the final survey

---

23 Pilot question: How has the feedback you have received as a part of the CMES process changed your work as an assistant principal?
questionnaire, this question was split into two questions to allow the participants to provide context for their answers. The thought behind the transition from qualitative to a mixed methods question for the last research question was to be able to categorize perception according to the quantitative question followed by getting more detail in the qualitative answer.

The pilot survey questionnaire had three open ended questions asking for feedback or comments at the end of the pilot survey. The answers given for these three questions contributed to the editing mentioned above. One of the questions also gave the researcher the ability to provide a range of time it would take to complete the survey to the potential participants.

For the web-based pilot survey questionnaire 27 assistant principals were asked to take the survey questionnaire. Sixteen of these assistant principals gave responses to the pilot survey. After an analysis of the results of the pilot survey, the researcher determined that certain changes needed to be made the survey instrument and how best to communicate the purpose of the survey instrument.

Reliability and validity. Reliability and validity “are ways of demonstrating and communicating the rigour of research processes and the trustworthiness of research findings” (Roberts et al., 2006, p. 41). Finding the validity of a research process is important because, “it is about the closeness of what we believe we are measuring to what we intended to measure” (Roberts et al., 2006, p. 41). Bush (2012) wrote of two types of validity: internal and external. Determining internal validity means that the research matches the intent of the research. External validity has to do with how the answers to the surveys by the sample can relate to the general population. Determining
the reliability of a research project, “describes how far a particular test, procedure or tool, such as a questionnaire, will produce similar results in different circumstances” (Roberts et al., 2006, p. 41).

Reliability. Using suggestions from Creswell (2012) and Bush (2012) a combination of tests for reliability were used for the main survey. Specific tests were used for the qualitative data. From these methods, the reliability of the instrument was determined.

Bush (2012) cited Youngman and suggested three ways in which to test reliability in survey research. The method used to determine reliability of the quantitative data for the survey instrument that will be used is to determine if the survey questionnaire is internally consistent. “If someone answers the questions a certain way at the beginning of the instrument and maintains those answers throughout, this is internal consistency” (Creswell, 2012, p. 161). Creswell (2012) recommended the use of both Spearman Brown formula and coefficient alpha to find the internal reliability of all questions within the instrument. The coefficient alpha is a “measure of the internal consistency of items on an instrument when the items are scored as continuous variables” (Creswell, 2012, p. 618). These continuous variables would be represented by the questions that have Likert Scale answers from which to choose. Coefficient alpha was used to determine if the instrument had internal consistency.

Bush (2012) also stated that reliability is impacted if questions are asked in the same way every time. The study sent the exact same survey questionnaire to all principals through an online format, therefore the impact of questions being asked in the same way was mitigated. This should prevent reliability from being an issue.
Credibility and transferability. Credibility and transferability were explained and determined for the qualitative part of the research for this study. Credibility (or internal validity) of qualitative research helped with determining how the research matched reality. Transferability (external validity) “is concerned with the extent to which the findings of one study can be applied to other situations” (Merriam, 2009, p. 223). In this study the method that would promote transferability in the study was maximum variation (Merriam, 2009).

According to Merriam (2009) one method that helps to ensure credibility of qualitative data is adequate engagement in data collection. For this strategy of data credibility “the best rule of thumb is that the data and emerging findings must feel saturated; that is, you begin to see or hear the same things over and over again, and no new information surfaces as you collect more data” (p. 219). This saturation point was reached for the data that was collected after the researcher reviewed the data multiple times and followed coding procedures to come to distinct categories. The researcher kept track of the number of times specific comments by the participants fit under specific categories. Those categories that had the most comments made by participants were the ones that were written in Chapter IV.

Merriam (2009) also stated that credibility of the researcher should be considered with qualitative data. To support researcher credibility, Merriam (2009) stated that a researcher should “look for data that support alternative explanations” (p. 219). Merriam (2009) further explained that some writers suggested looking for data within findings that might be contrary to the emerging findings. In this study, when positive perceptions were
asked of participants, some participants still gave negative answers. These findings are reported and analyzed in Chapter IV.

The strategy used for attempting to enhance transferability in this study was maximum variation. Using this strategy, the researcher tried to seek “variation or diversity in sample selection to allow for a greater range of application of the findings by consumers of the research” (Merriam, 2009, p. 229). For this study, the researcher attempted to send a survey questionnaire to all principals in the state of Colorado who were being evaluated using the CMES. The demographics of the participants show participants from many different demographics. Though the principals selected to be a part of this study had to have experienced receiving feedback as a part of the CMES process, there were 164 school districts in the state of Colorado who currently use CMES for principals. These 164 school districts represent all demographics that were a part of this study. The attempt to enhance transferability for this study through maximum variation should have allowed for a greater variation in application for those who might read this study (Merriam, 2009).

*Triangulation.* Triangulation is “essentially a means of cross-checking data to establish its validity” (Bush, 2012, p. 84). The method of triangulation used for this study was a methodological triangulation. Bush (2012) stated that this form of triangulation is “where strategies or methods are mixed to corroborate against one another” (p. 85). This study is a mixed-methods study where the researcher compared some of the qualitative and quantitative results, and consolidated quantitative results according to how participants answered the third question. This mixed-methods study allowed for a methodological triangulation of data.
Collecting data using the survey questionnaire. The pilot study was completed, and edits were made prior to completing the final survey. The researcher followed a specific process to send and collect the survey questionnaire to the principals being evaluated using the CMES. This process spanned six weeks. The following paragraphs will explain the data collection process including: suggested methods to ensure a strong response rate and how the survey was disseminated and collected.

Prior to explaining how the researcher tried to ensure a strong response rate for this survey questionnaire, expectations for response rates for surveys was researched and detailed below. Sivo, Saunders, Chang, and Jiang (2006) conducted a meta-analysis of survey research and concluded, “the average response rate ranged from 22% to 59.4%” (p. 356). Baruch and Holtom (2008) studied response rates of survey questionnaires from 1975, 1985, 1995, 2000, and 2005 and found that the mean response rates in survey research dropped from 64.4 percent in 1975 to 48.3 percent in 2005.24 In 2005 the range in response rates for survey questionnaires given to individuals (157 studies) was from 3 percent to 91 percent with a Standard Deviation of 21.2. With the research studies by Sivo et al. (2006) Baruch and Holtom (2008) in mind, a goal of over 50 percent response rate was set by the researcher for this study.

The researcher also did a non-respondent survey by email of principals who did not respond to the survey. Qualtrics tracks those who have not opened or started the survey when sent as a link to their email address (Qualtrics, October 22, 2014). It shows the number of people who have responded to the survey, completed the survey, and/or

24 When comparing the years 2000 and 2005 Baruch and Holton (2008) believed that the response rates seemed to be stabilizing over this period of time.
did not open the survey. The researcher found 15 email addresses who did not respond to the survey or started the survey without responding to even the first question (the opportunity to opt out). The researcher then emailed these non-participants asking for them to take an edited (shortened) version of the survey questionnaire. A follow up email was sent to this group.

Creswell (2012) gave a few suggestions to ensure a strong response rate. These suggestions include: studying a problem of interest for the respondents, use a brief instrument, pre-notify the respondents, and use good follow up procedures. These suggestions all were used to strengthen the probability of a strong response rate.

This survey should have interested the assistant principals who are respondents to the pilot survey, and the principals who will be the respondents of the main survey. Evaluation is important to these two groups because of its summative purpose of retention or termination for obvious reasons: employment. This survey should also capture the interest of principals and assistant principals because the purpose of feedback using the CMES should be to impact the practice of these two groups (State Council for Educator Effectiveness, 2011). The evaluation tool and system should have been top-of-mind for these principals. “If individuals in the sample are interested in the issue, they will be more apt to complete the survey” (Creswell, 2012, p. 391). Therefore, a survey asking questions about the feedback received using this evaluation tool should be of interest to both the principals and assistant principals being evaluated using the CMES system.

The second suggestion by Creswell (2012) was to use a brief instrument. The researcher created a survey questionnaire that did not have any redundant questions, and
asked questions referencing the feedback received by those being evaluated using the CMES. One of the follow up questions for the pilot study was how long the survey took the respondents. The most frequent time that the pilot survey took the respondents was 10 minutes. The range of time the pilot study took the respondents was from 8 minutes to 18 minutes. This was a brief instrument.

Creswell (2012) also recommended pre-notifying the survey respondents that they will be receiving a survey, and after the survey is sent, to have good follow-up procedures. An email was sent to all potential respondents two weeks prior to sending the survey. This email had an explanation of what was to come, also a copy of the Consent Form for Human Participants that was approved by the Institutional Review Board (IRB) at the University of Northern Colorado (UNC).25 This email was then followed up by an email that included the actual survey instrument. This email gave a time-frame the survey instrument was open. Two weeks after the survey questionnaire was sent out, a second email was sent to either remind all survey participants to respond, or to those who have not responded to remind them to take the survey. A final email was sent to the same group as the previous email in order to give them one final reminder.

Sending a survey questionnaire that is of high interest to respondents and having a procedure to send the survey instrument that appropriately communicates the survey instrument to the respondents are both important to creating a strong survey instrument (Muijs, 2012; Creswell, 2012). The researcher was hopeful that the format of the survey and the chronology of the survey dissemination would ensure a strong response rate. Besides the process used to create and distribute the survey questionnaire, Creswell

---

25 The approved Consent Form for Human Participants can be found in Appendix E.
(2012) suggested two other valuable points to take into consideration when collecting data.

It is important to make sure the instrument and process used when collecting data is standardized (Creswell, 2012). “If procedures vary, bias can be introduced into the research” (p. 169). This survey questionnaire was standardized because the same instrument was sent using Qualtrics online survey software. Everyone who takes the instrument received the same reminders in the same format. The participants also received the same survey instrument in the same format.

A second aspect of data collection that Creswell (2012) emphasized is the potential ethical issues that come about when collecting data. Creswell (2012) first underscored the importance of protecting the anonymity of the survey respondents. Creswell (2012) also highlighted that the information given by the survey participants cannot be shared with anyone not involved in the project. Creswell (2012) finally stated that the participants must be given an option to opt out of the research even if they initially agree with participating.

All the ethical issues tackled by Creswell (2012) was addressed when a research application was sent to and approved by the IRB at UNC. Qualtrics, through their technology, ensures the confidentiality of the survey questionnaire participants. Besides the researcher, the only people who saw the survey research were members of the research committee for this project and a couple members of the Research Consulting Lab at the University of Colorado. Finally, the survey questionnaire participants will be allowed to opt out at any point. This will be explained in the consent for human

26 The UNC approval of the IRB can be found in Appendix F.
participation that was approved by the IRB through UNC, the first question of the survey also had this explanation; finally, the emails sent prior to and after the survey questionnaire included this information.

In conclusion, data collection for this research project occurred using the *Qualtrics* software that was supplied to the researcher by staff of the University of Northern Colorado. To ensure a strong response rate, a brief survey questionnaire was made, which was distributed using suggested methods by Creswell (2012). To attend to potential ethical issues and standardization concerns, the survey questionnaire was not distributed until the IRB at UNC approved the research process, and the approved survey questionnaire was sent to all potential participants. The data was collected and stored in the *Qualtrics* system. After data collection concerns have been addressed, the data analysis procedures will be explained.

**Data analysis.** The next step after the research has been collected is to analyze the data. According to Hibberts and Johnson (2012) the type of data analysis to be used for this study will be “multidata – multianalysis” where both quantitative and qualitative data will be used and these data will be analyzed using quantitative and qualitative methods (pp. 135-136). The researcher first analyzed the quantitative data. The qualitative data was the second set of data that was analyzed. In this section, the quantitative analysis will be explained first. Second, will be the type of qualitative analysis. The final part of this section will be a clarification of the convergent design analysis in relation to this specific study. Figure 4.1 shows the data analysis process.
Figure 4.1. Data analysis process for this mixed methods study.

**Quantitative data analysis.** The first set of data analyzed was the quantitative data. According to Dixon and Woolner (2012), to properly analyze the data, the following process should be used. First, the identified variables within the survey questionnaire were classified as either nominal or ordinal variables. Then the data was coded to “turn it from words or ideas to numeric form” (Dixon & Woolner, 2012, p. 344). The first set of statistics that were analyzed were the descriptive statistics. After the descriptive statistics were analyzed, the inferential statistics were evaluated. The description of the results of this analysis will occur in Chapter IV and Chapter V.

To analyze the quantitative data, Statistical Package for the Social Sciences (SPSS) software was used. To be able to download the information, the first step that was followed was identifying which type of variables were used in order to assign them numbers and appropriate quantities. The two types of variables that were identified are nominal variables and ordinal variables. Nominal variables are variables based on classifications or groupings (Dixon & Woolner, 2012). Dixon and Woolner also stated that for these types of variables the number assigned to each of the options or choices do not equal an actual quantity. The nominal variables for this study were both the demographics questions and the list answer question, which asked principals the amount
of feedback they received from their evaluators. The second set of variables identified on this survey questionnaire were the ordinal variables. According to Dixon and Woolner (2012) ordinal variables have values that can be thought of as a ladder or a staircase, where ‘4’ is consistently above ‘3’, but “there is no certainty that the steps are equal distances apart” (p. 342). These types of questions have answers that might be a rating scale from strongly agree to strongly disagree that “might be measuring opinion in some sense but it is not like a ruler or a tape measure” (Dixon & Woolner, 2012, p. 342). In this specific survey questionnaire, the questions with Likert Scale type answers would be considered ordinal variables.

After designating the types of variables that each answer represented, the next step is to code the data to, “turn it from words or ideas to numeric form” (Dixon & Woolner, 2012, p.344). Answers to each of the questions were assigned a number (by SPSS). Dixon and Woolner (2012) recommended for those questions that have the Likert Scale type answers, that all answers are consistent, and therefore that all numbering/coding for these answers be consistent also. After assigning appropriate numbers for each of the answers, the next step in data analysis is inputting the information from the survey questionnaires into the SPSS system. For this research project, the Qualtrics software had a simple process to enter the information into the SPSS software.

Descriptive statistics. After the information was downloaded into the SPSS software, the first set of data analysis was to determine the descriptive statistics. The initial descriptive statistics to be analyzed was univariate analysis. After the univariate analysis the questions were then looked at in comparison to one another to see if there
were any trends in answers. From the descriptive statistics came an understanding of the inconsistencies in the data and the general trends in answers from all of the participants.

Dixon and Woolner (2012) suggested to analyze the answers to the survey questionnaire through univariate analysis first. This is to determine errors in the data and to take an initial look at the data. The three types of descriptive statistics that were used to analyze the survey questionnaire data were the frequency distributions, finding measures of central tendency, and using measures of dispersion. From these data should come an understanding of any potential mistakes in the survey questionnaire, and an understanding of the trends in answers for the principals as a whole.

When using frequency distributions for nominal and ordinal data, researchers use bar charts or histograms to see the data (Dixon & Woolner, 2012). These charts “will show the number of times the variable values occur in your data set” (Dixon & Woolner, 2012, p. 346). The initial check of the data showed if there were any mistakes when inputting the data, and whether there were any answers that might be outliers. When looking at the frequency distributions, the researcher could also find whether the answers given are skewed positively or negatively, or if the answers follow a normal curve. From the frequency distributions came a set of data that gives a broad perspective of principals’ perceptions of the evaluation system. For this study, bar charts were referenced, but tables showing actual numbers were used.

After creating frequency distributions, Dixon and Woolner (2012) recommended looking at the measures of central tendency when analyzing the data. The measures of

---

27 Dixon and Woolner (2012) stated that it is best to use bar charts for nominal or ordinal data.
central tendency are the mean, median, and mode of each set of answers or variables. The measures of central tendency can help to describe the trends in answers given by principals and would be used in the calculation of inferential statistics. The mean is the total of the scores divided by the number of the scores. It provides the average score of all the statistics combined. For the descriptive statistics, the mean was only used for the demographics in this data analysis. The median is the middle of a set of scores. The mode is the most frequently occurring score in a set of scores. The mode is most useful when describing categorical variables (Dixon & Woolner, 2012).

A cross-tabulation of data was also done for the final analysis. This cross-tabulation considered the participants’ perceptions in comparison to the frequency of feedback reported by the participants. The reason behind this cross-tabulation was to see if there might be any relationship between attitudes of principals and the amount of feedback they were receiving.

By using descriptive statistics to analyze the data there is an opportunity for a researcher to start the research analysis process. Frequency distributions provide an opportunity for the researcher to create a visual of the answers for each of the individual questions. Measures of central tendency allow the researcher to measure the average, the central number in a set, or the numbers that appear most frequently. A cross-tabulation of data compares the perceptions of principals with the reported practice of their evaluators. From these sets of answers will come a comprehensive understanding of the trends of answers.

_Inferential statistics_. Inferential statistics was the next way in which data from this survey were analyzed. The reason behind using inferential statistics is to be able to
compare two or more groups or two or more variables (Creswell, 2012). “The basic idea is to look at the scores from a sample and use the results to draw inferences or make predictions about the population” (Creswell, 2012, p. 187). When comparing the relationship between research questions and demographics with two options for responses, the chi-square test was used. The chi-square test “is used to determine whether frequency counts are distributed identically across different populations” (Stat Trek, 2016, para. 1). When comparing the relationship between each research question and demographics with multiple options a univariate analysis of variance was used (ANOVA). According to Creswell (2012) an ANOVA is a statistics test frequently used in educational research for group comparisons with one or more independent variables and one dependent variable. When the researcher worked with a consultant from the University of Northern Colorado’s research lab, this test was used to compare demographics groups’ experiences with the research questions.

Chi-square test was used for this survey in two ways. The first was to compare whether or not the answers of the participants for the three research questions were different according to the gender of the participant. The second way that chi-squared test was used was to determine if participants from one region of the state answered questions differently than all other parts of Colorado.

The univariate ANOVA was used to see if there were any significant differences according to categories of demographics when the participants answered the three research questions. The demographics for which an ANOVA was run were: free and reduced meals rate of the school, school size, principal total years, and principal years in current position. These demographics were chosen because the expectation of the
researcher was that there might be different in how each of these groups might answer the research questions.

- Free and Reduced Meals rate of a school – Support provided by a principal evaluator might differ according to the amount of low income students in the school.
- School size – The experience with CMES/ feedback might differ according to school size because size of school might impact the amount of support.
- Total years as a principal – The more experience of a principal might mean that a principal may require less feedback.
- Years in current position – Principals who might be less experienced in their current position might receive more feedback.

Each of these demographics were thought to potentially provide some discrepancies when compared to others within different demographics.

To calculate the chi-squared test and the ANOVA, some things must be established both before and while calculating (Dixon & Woolner, 2012; Creswell, 2012). Initially, the null hypothesis and the alternative hypothesis must be established. The null hypothesis means that the difference in means between to two variables, that which are being compared, is zero. The alternative hypothesis would be anything more than zero. To determine the alternative hypothesis, the alpha level must be set. The alpha level “reflects the probability level that reflects the maximum risk you are willing to take that any observed differences are due to chance” (Creswell, 2012, p. 188). For this study the alpha level was set at .05. If the number calculated using a chi-square test or the
ANOVA is less than the set percentage, it is proof that the difference in means between the two variables is statistically significant. If the opposite is true, then the means of these variables is not statistically significant (Dixon & Woolner, 2012; Creswell, 2012).

For the answers to this survey questionnaire, inferential statistics were used to see if there was any difference between how participants responded from different demographics. These potential differences were calculated using the chi-squared test and univariate ANOVA. If the calculation showed to be statistically significant, the specific demographics or questions used could be shown to have different perceptions of the answer to that specific question.

**Conclusion of quantitative data analysis.** Descriptive and inferential statistics will be used to analyze the quantitative data of this survey questionnaire. The descriptive statistics will help the researcher develop a general understanding of the research through frequency distributions, and measures of central tendency. By using both sets of statistics to analyze the data, a comprehensive picture of the perceptions of principals of the Colorado Model Evaluation System and the feedback received through this system should be found.

**Qualitative data analysis.** For the qualitative data analysis, the first thing that must be understood by the researcher was the nature of the central phenomenon which is the main topic to be studied (Creswell, 2012). In this study, the central phenomenon was the principals' experience with the purpose of the Colorado Model Evaluation System for Principals, which is essentially to provide formative feedback so principals might improve their practice (State Council for Educator Effectiveness, 2011). The next step in the qualitative part of research study was to identify the type of research to be completed.
The qualitative part of this study was a basic qualitative study. Merriam has found that this type of qualitative analysis is probably the most common in applied fields such as education (Merriam, 2009). To describe that which makes a qualitative study a basic qualitative study, Merriam (2009) stated that researchers “would be interested in (1) how people interpret their experiences, (2) how they construct their worlds, and (3) what meaning they attribute to their experiences” (p. 23). For this study, the researcher was looking for how principals interpreted their experience with feedback. The participants in this study were also asked how feedback was impacting their practice, thus asking them how they constructed their job from the feedback they received. All the qualitative questions followed closed ended quantitative questions to allow participants to provide context for their answers and their experiences, therefore providing meaning to their experiences. Merriam (2009) stated that the primary goal of a basic qualitative study is to uncover and interpret how people construct meaning in their lives. In this study the participants were asked to construct the meaning of feedback received as a part of the CMES. Through the qualitative answers, the researcher uncovered and interpreted these meanings.

After all the qualitative data were collected by survey questionnaire, they were analyzed. First, the researcher organized the data for analysis (Creswell, 2012). *Qualtrics* facilitated this through their website. After closing the survey, the researcher downloaded a *Comma Separated Value (CSV)* file that can then be saved as a *Microsoft Excel Document*.\(^{28}\) These data were then uploaded into *NVivo* and printed before starting the analysis.

\(^{28}\) This will be how the quantitative data was downloaded and stored also.
After organizing the data, the researcher then went through a series of steps to analyze and categorize the qualitative data. The first step was to get a general feel of the data. The second and third steps were methods of coding as described by Saldana (2016) as first cycle and second cycle coding. The steps followed for this research were:

1. Open coding
2. First cycle coding (Saldana, 2016)
   a. In Vivo Coding
   b. Descriptive Coding
   c. Magnitude Coding
3. Second cycle coding (Saldana, 2016)
   a. Axial Coding

According to Merriam (2009) open coding occurs when a researcher is open to any finding while initially reviewing the data. For the analysis of this research, the open coding was done question by question. Each research question was organized in NVivo and then printed so the researcher might mark on the paper. During the initial read of the data aligned with each question, the researcher would mark or make note of things noticed in the answers for each question.

After open coding, the researcher then used in vivo coding to determine the initial codes. According to Saldana (2016), open coding “refers to a word or short phrase from the actual language found in the actual language found in the qualitative data record” (p. 105). Saldana (2016) also stated that the voice of the participants is more likely communicated when their exact words are used or coded. For this particular study in

---

29 The questions can be found in Appendix D: Q15, Q16, Q17, Q18, Q20, Q21, & Q22
vivo was used to see which words were the most common to show up when analyze the data from all participants. When particular words or phrases appeared in comments made by the participants, these words were color coded.

A second method of coding used in conjunction with in vivo coding was descriptive coding. “Descriptive Coding summarizes in a word or phrase the basic topic of a passage of qualitative data” (Saldana, 2016, p. 102). After reading through and color coding according to exact words and phrases, those that did not have exact words to be coded, were then given a descriptive code according to the topic that was described by the participants.

A third method of coding that was used for the third research question was Magnitude Coding. Magnitude Coding “applies alphanumeric or symbolic codes and/or subcodes to data to describe their variable characteristics such as intensity or frequency” (Saldana, 2016, p. 82). These magnitudes were assigned according to participants’ Likert responses to the quantitative questions. These participants were grouped according to positive, neutral, and negative responses. The researcher then went through open coding, In Vivo coding, and Descriptive Coding for each of these categorized data.

At the end of the first cycle of coding for each of the questions, the researcher then lumped the codes together to condense them further for analysis. The initial lumping of the codes happened for each question. Then Q15 and Q17 were combined because the intent behind the questions were to receive positive responses. Q18 and Q20 were then combined because the intent behind asking these questions was to find opportunities to improve. The second cycle coding methods used in this study are described below.
The method of coding used for the second cycle of coding was axial coding. Saldana (2016) described this type of coding as a way of reorganizing data under some dominant codes and to determine (and report) dominant codes and not reporting the less important codes. Dominant codes are to be organized into categories. Dominant categories may have subcategories that are connected to the categories. Saldana (2016) also stated that during axial coding that a researcher should want to achieve saturation. Saturation occurs when no new information is found when analyzing the data. For the research in this dissertation, axial coding was used as the method to create the final organization of the qualitative data.

After codes were created, the researcher then created categories of codes, which all demonstrated certain themes. These themes were then further analyzed and major categories were created that connected certain categories to one another. Those categories that only had a small number of participants’ comments attached to them were either connected to an appropriate major category or not reported. Major categories were then given definitions by the researcher to communicate the connections between all subcategories, and ultimately, connecting all comments by participants that fit under each major category.

In Chapter IV, the number of times a major category was mentioned by participants was also reported. This was to create a connection to the magnitude of each of the major categories. Every time a comment by a participant fit under a major category, this was counted by the researcher of this study. In Chapter IV, when it is

30 Major categories was one of the ways in which Saldana (2016) described the categories that were consolidated or refined after multiple rounds of coding.
written that a major category was mentioned by participants a certain number of times, this was a result of axial coding where the participants’ comments have a purposeful connection to the major categories.

**Mixed methods analysis.** As was stated at the beginning of this section, the study followed up the quantitative and qualitative analysis by doing a mixed analysis with the data. The method of analysis used was identified by Creswell (2012) as convergent design analysis. When the researcher converged the data for this study, initially the results of the two parts of the survey questionnaire were placed next to each other to create a discussion about both results, thus comparing the results (Creswell, 2012). The second part of the convergent design analysis was to dive deeper into the relationship between both sets of data by directly comparing the quantitative and qualitative results, thus consolidating data (Creswell, 2012).

When the two pieces of data were placed next to each other, the initial step was to compare the results of the descriptive statistics from the Likert Scale answers to the trends in information for those questions that follow the questions with Likert Scale answers. All of the qualitative questions in this survey were follow up questions to the quantitative questions and provided a natural way to compare the quantitative themes from the descriptive statistics (median, mode) and the themes that can be found for each of the answers. This was the initial step followed when mixing the research results.

The next step in the mixed methods analysis was to look further into the answers of the quantitative questions and comparisons to the themes of the qualitative data. Initially, each of the quantitative questions were studied first. For the first two questions, when principals reacted (positively, neutral, negatively) while answering a specific
question, what were the general themes of the qualitative data that were pulled from this pool of people? For the third question: the analysis used was convergent design: consolidating results. The answers given by the sample principals to the qualitative follow up question were sorted according to whether or not the principal answered the quantitative question with Likert responses positively, neutral, or negatively. The qualitative responses were then analyzed according to these re-categorized data.

The purpose of this final piece of analysis was to allow the strengths of qualitative and quantitative data collection and analysis to be highlighted. According to Hibberts and Johnson (2012) a strength of quantitative research is “it is useful when studying large groups of people and providing the etic perspective (and) it produces standardized measures of relationship” (p. 123). Qualitative research “can explicate insiders’ perspectives with authenticity (and) idiographic causation (i.e. causation that we observe, purposively produce or experience in our lives) can be documented” (Hibberts & Johnson, 2012, p. 124). With both sets of strengths in mind, a mixed methods study was created.

**Researcher Bias**

The following paragraphs identify bias by the researcher and outlines the steps that were taken to prevent this bias from impacting the reported results. Schwandt (2001) stated that bias “means individual preferences, predispositions, or predilections that prevent neutrality and objectivity” (p. 15). This researcher’s experiences with the Colorado Model Evaluation System has created a set of biases that will be explained. Steps were taken to mitigate these biases. This researcher’s individual preferences as far as feedback given as a part of an evaluation system, and whether or not principals are
receiving consistent feedback that is impactful, tie into the biases the researcher has because of his own evaluation experience. The way of getting/giving feedback that is preferred by this researcher is Cognitive Coaching which is not the method used by any of this researcher’s evaluators since working in administration (until recently). This researcher also has not received consistent feedback that has anything to do with performance. These preferences then tie into the predisposition of the researcher while being evaluated within the CMES system. The predisposition of this researcher towards this system has been impacted by the researcher’s lack of positive experience with the CMES. The CMES has been handled as an additional task and has not been coordinated with this researcher’s actual practice. The final piece of bias that this researcher has towards CMES is the predilection of this researcher towards using this system as a formative tool to improve practice as it is intended (State Council for Educator Effectiveness, 2011). Because of the inconsistencies of feedback and the experience by the researcher of feedback that does not necessarily positively impact practice, this researcher has expectations of feedback in this system that have not been fulfilled at this moment. Though these biases exist in the researcher, they were mitigated through the following efforts.

To avoid bias from entering into this research paper, the researcher wrote and reported the research findings in an ethical way (Creswell, 2012). After ensuring that the quantitative data was valid and reliable and the qualitative data was credible and transferable, the data was uploaded to SPSS and NVivo in order to appropriately analyze the data. Using these two programs to do the initial sorting of information prevented the researcher from allowing bias to enter the initial sort of information. The advisor to the
researcher was also be asked to help the researcher process the data as a part of writing Chapter IV and Chapter V. Finally, the researcher kept his own bias at the top of mind when analyzing the data, not to promote the bias, but to make sure the bias does not make its way into the research report.

**Assumptions of This Study**

There are certain assumptions made by the researcher. The first assumption was that because feedback is supposed to be a valuable part of this evaluation system, that principals were actually receiving feedback from their evaluators. The second assumption was that these evaluations were being used for their formative and summative purposes; not just for the summative hiring and firing process. Certain assumptions were also made about how the survey questionnaire was sent to principals. Since the *Qualtrics* system requires the use of a computer (or mobile device) in order to answer the survey questions, all principals should have had access to computers in order to answer the survey questionnaire. Another assumption is that Qualtrics’ software available for the assistant principals and principals to access the survey using their phones did not have any complications. Another assumption made about the sending of the survey questionnaire is that *Qualtrics* has the technology and the know-how to build survey technology that could avoid being sent to spam. With district firewalls set at different levels, this could have created some difficulty. A final assumption made by the researcher had to do with the email list purchased by emaillistus.com. It was assumed by the researcher that all of the information purchased was accurate and up-to-date.
Conclusion

The Colorado Model Evaluation System for principals is supposed to be used as a summative and formative tool for evaluation. This research focused on the evaluative feedback received by Colorado principals and their perceptions of the impact this feedback has on their practice. The hope is that the information provided by this study will provide a foundation on which other studies would be based and potentially help with the continuous improvement of the CMES. The next logical step for the researcher would be a second study based on any questions or topics that arose from this research built using the research questions:

Q1 What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process?

Q2 What are principals’ perceptions of the CMES evaluation feedback?

Q3 According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

The results of this study are discussed in Chapter IV and Chapter V.
CHAPTER IV

ANALYSIS

The purpose of this study was to understand the perceptions of feedback by Colorado principals who were being evaluated using the same evaluation system, namely the Colorado Model Evaluation System (CMES). The participants’ perceptions were gathered using a survey questionnaire and were analyzed using a mixed methodology. Although there were 24 questions on the survey questionnaire, the (non-demographics) quantitative and qualitative questions were all crafted to help collect the necessary data required to answer the following research questions:

Q1 What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process?

Q2 What are principals’ perceptions of the CMES evaluation feedback?

Q3 According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

Following this introduction there are six sections to this chapter. First is a description of the response rate for this survey questionnaire. Second is the calculation of internal consistency of the three research questions with Likert scale responses. Third is the description of the survey questionnaire participants. The Fourth and fifth section will be the quantitative and qualitative results to the first research question. For each of these sections, the researcher explained the quantitative results, followed by the qualitative themes, and concluded with a comparison between the quantitative and qualitative results. In the sixth section the researcher will explain the quantitative results first. The
second half of the research analysis for research question #3 was a mixed methods analysis. In the final section of this chapter, the researcher will give an overall synopsis of the research results.

Survey Response and Demographics Analysis

In the following pages, the participants are profiled based on response rates and their answers to the demographic items on the survey. The numbers for the response rate were detailed first. The demographics of the participants followed the response rate. When the categories for demographics had comparable data, the sample was compared against data provided by the Colorado Department of Education.

Response Rate for this Study

The goal for response rate of this survey was over 50 percent and the actual response rate was not close to this intended response rate (16.02%). This response rate was based on 949 principals who were sent the final survey. An explanation of the final response rate is detailed below.

How the final response rate was determined. The researcher sent an email from his University of Northern Colorado email account.31 This initial email was sent to a list of 1604 principals in the state of Colorado. The researcher then went back to the principal list and edited out email addresses that fit under the following categories: cease and desist school districts, email addresses of principals who were not being evaluated using CMES, and individual requests to not be a part of the survey. The final survey was then sent to 949 principals’ email addresses. After closing the survey, there were 187

31 The initial email is Appendix G
principals who responded in some manner to the final survey. Eighteen of the principals responded that they did not wish to take the survey. Thirteen responded that they had not been evaluated using CMES for principals the previous year and were removed. Four principals were removed because they answered the demographics questions and nothing else on the survey. The total number of participants whose data were used for the final analysis was 152 or 16.02 percent.

**Potential reasons behind response rate.** The researcher had eight potential reasons behind the low response rate:

- **Cease and Desist Orders:** When the researcher first sent the survey questionnaires to the email list purchased from emaillistus.com two school districts responded with cease and desist letters stating that these school districts did not authorize research unless it went through its IRB process. The researcher decided not to proceed with both of these processes because neither process was timely for this research study. These school districts were excluded from this study despite the fact that this decision meant a large number of principals, who may have completed the survey, were denied the opportunity to participate. Their participation could have positively impacted the overall response rate.

- **Inaccurate Information:** As the researcher went through the email list to delete emails because of the cease and desist letters, he realized that the list had inaccuracies. The extent of these inaccuracies was not known by the researcher, but these inaccuracies had a negative impact on the overall response rate.
• Duel Role: Many of the small school districts in the state of Colorado have principals who are also superintendents and might be evaluated by school boards instead of using this system.

• Not the Target Population: Although the researcher did much pre-sorting and deleting of email addresses of principals who were not a part of the CMES, there was a chance that some of the principals that were invited to complete the survey ignored it because they were not evaluated using CMES.

• Email Overload: Principals typically receive a multitude of emails each day and some messages can get buried or overlooked, even when the recipient may have been willing to participate in the study.

• Survey Questionnaire Fatigue: In a study on response rates, Baruch and Holtom (2008) stated that certain professions get over surveyed and “the result is a large number of target individuals or firms who are fatigued and therefore refuse to respond to non-essential questionnaires” (p. 1142). Principals receive many requests asking for information that help to inform researcher, companies, and other organizations. This might have impacted the response rate to this survey questionnaire.

• Time Commitment: Finally, taking 10 to 20 minutes to complete a survey might be too much time for a principal to work on outside of the daily responsibilities.

The researcher also tried to reach out to those who did not complete the survey in order to look into reliability and validity. A final request was sent to fifteen principals on
the email list who did not respond during the original survey window. One of these principals responded to this request, but stopped responding after a request for interview and/or request to respond to follow up emails.

The response rate did match the range in responses when looking at Baruch and Holtom (2008) found when analyzing response rates for individuals taking surveys in 2005 (3% - 91%). It did not match the range found by Sivo et al. (2006) for survey research (22% - 59.4%). The final response rate also did not match the intended response rate proposed by the researcher (50 percent). Although the researcher is not statistically comfortable with the final response rate, the researcher was comfortable to move forward with the data analysis. First, the email list used might have included dated email addresses, thus the initial number of emails sent might have been less if the list was up-to-date. Second, being in an occupation that might have survey fatigue might impact the likelihood of responses from non-essential organizations (such as someone writing a dissertation). Third, though it did not impact the response rate of the 949 principals who were sent the survey questionnaire, the researcher was sent “cease and desist” communication from two extremely large school districts. These school district processes would have required a delay in data collection that the researcher determined was too long to wait for data collection. The researcher had to remove 655 email addresses from the original email list to abide by these school districts’ requests.

The researcher also believes that sufficient effort was put into soliciting the responses. An original email was sent from the researcher’s University of Northern Colorado email account. This email was followed by the survey, which was open for 29 days. All of participants who did not complete the survey received two follow up emails
prior to the email distribution being closed. Overall, potential participants could have received up to four requests to complete the survey during the survey questionnaire window.

Finally, the researcher tried to enhance transferability through what Merriam (2009) calls maximum variation when it came to the participant sample by sending the survey questionnaire to principals from all over the state of Colorado who were being evaluated using CMES (Merriam, 2009). The demographics of the participants showed that they represented all subgroups in every demographics. Though the response rate was less than the researcher had hoped for, the comfort level with analyzing the data is pretty high because of the above-stated reasons.

**Internal Reliability of the Instrument**

As was suggested by Creswell (2012), the author chose Cronbach’s Alpha to determine the internal reliability of the 152 responses that were determined using the above-mentioned reasons. There were three main questions in the survey that provided a majority of the data needed to answer the research questions. These three items were: 1) What are the principals’ perceptions of the CMES evaluation process?, 2) What are the principals’ perceptions of the CMES evaluation feedback?, 3) According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principal’s practice?. The Likert responses of the three main research questions were entered into Statistical Packages for the Social Sciences (SPSS). Cronbach’s Alpha came back as
.683 for these three items. The internal consistency of these three items came back with a high Cronbach’s Alpha.

**Survey Participants**

In this section of the chapter, the demographics information of the survey questionnaire participants are presented. The explanation of the demographics will go in the following order: gender of the principal, gender of the evaluator, type of evaluator, years as principal, years in current position, type of district, location of district, free and reduced meals rate, size of school, type of school, and final effectiveness rating. CDE data was included for those demographics that are collected by CDE and were collected by the researcher of this study.

**Gender of principals.** For this survey 68 (44.7%) of the participants were male and 84 (55.3%) were female. The Colorado Department of Education has data that combines both the total numbers of principals and assistant principals and sorts them by gender. The male population of both principals and assistant principals is 43.92 percent (CDE, January, 2017). The female population of both principals and assistant principals is 56.08 percent (CDE, January, 2017).

**Evaluator demographics.** The reported gender of the evaluators of these principals was 82 (53.9%) male and 70 (46.1%) female. The type of evaluators for these principals were 74 (48.7%) superintendent, 49 (32.2%) director of secondary or 32

Although the standards for what makes a “good” $\alpha$ coefficient are entirely arbitrary and depend on the theoretical knowledge of the scale in question, many methodologists recommend a minimum $\alpha$ coefficient between 0.65 and 0.8 (or higher in many cases); $\alpha$ coefficients that are less than 0.5 are usually unacceptable (Goforth, 2015).
elementary education, 14 (9.2%) assistant superintendent, and 15 (9.9%) other designations.

**Years as a principal.** The number of years the participants have been principals varied greatly. The mean number of years the participants have been principals was 8.28 years with a range of one year to 27 years. The median years of these participants was 7 years. The most common number of years (mode) by these principals was 5 years with 21 participants answering this number. The year with the second largest number was 4 with 15 participants answering with this number. A general breakdown of years as a principal is highlighted in Table 4.1.

Table 4.1

<table>
<thead>
<tr>
<th>Range of Years</th>
<th>Number of Participants (n=152)</th>
<th>Percent of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>64</td>
<td>42.10</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>42</td>
<td>27.63</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>28</td>
<td>18.42</td>
</tr>
<tr>
<td>16-20 Years</td>
<td>12</td>
<td>7.89</td>
</tr>
<tr>
<td>21-27 Years</td>
<td>6</td>
<td>3.95</td>
</tr>
</tbody>
</table>

**Years as principal in current position.** The participants were also asked how many years they had been in their current position. One hundred and forty-nine respondents gave number responses to this. One respondent stated “principal” and other stated “MS principal” and a third left this question blank. The mean number of years for these principals was 5.16. The median number of years was 4.0 years. The mode for

---

33 A request was made of CDE for any demographics information that relates to this study. This was one category that CDE did not respond with any information.
these principals is 4.0 years also. The breakdown for these principals is on the next page in Table 4.2.

Table 4.2

<table>
<thead>
<tr>
<th>Years in position</th>
<th>Number of Participants</th>
<th>Percent of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5 Years</td>
<td>100</td>
<td>67.11</td>
</tr>
<tr>
<td>6-10 Years</td>
<td>33</td>
<td>22.15</td>
</tr>
<tr>
<td>11-15 Years</td>
<td>14</td>
<td>9.40</td>
</tr>
<tr>
<td>16-20 Years</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21-27 Years</td>
<td>3</td>
<td>2.01</td>
</tr>
</tbody>
</table>

**Location of school district.** The location of the school district was the next demographic to be asked of the participants. For this question 151 participants gave responses. One participant answered “other” but did not explain where their district was. The largest number and percentage of principals to respond to this question were from the Front Range. The total numbers and percentages can be found in Table 4.3.

Table 4.3

<table>
<thead>
<tr>
<th>Location of School District</th>
<th>Number of Participants</th>
<th>Percent of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Mountains</td>
<td>6</td>
<td>3.97</td>
</tr>
<tr>
<td>Central Mountains</td>
<td>8</td>
<td>5.30</td>
</tr>
<tr>
<td>Southern Mountains</td>
<td>15</td>
<td>9.93</td>
</tr>
<tr>
<td>Front Range</td>
<td>70</td>
<td>46.36</td>
</tr>
<tr>
<td>Western Slope</td>
<td>27</td>
<td>17.88</td>
</tr>
<tr>
<td>Eastern Plains</td>
<td>33</td>
<td>21.85</td>
</tr>
<tr>
<td>Other (No response)</td>
<td>1</td>
<td>0.66</td>
</tr>
</tbody>
</table>

**Type of school district.** This demographic had 150 participants respond to the question. The largest number of survey participants were from was rural school districts with a total of 83 (54.60%) participants. This category required a consolidation of categories for the sample and for the state data. This was to match the categories as
accurately as possible. Example - The reason why rural and outlying town were
combined was because when the researcher considered the districts that were labeled
“Outlying Town” those outlying towns seemed to be in Rural areas.

Table 4.4

*Type of school district: State of Colorado verses participants*

<table>
<thead>
<tr>
<th>Type of District</th>
<th>Participant Percent (Principals)</th>
<th>Colorado State Percent (Districts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban/ Suburban (Denver Metro/ Urban-Suburban/Outlying City)</td>
<td>45.40</td>
<td>25.14</td>
</tr>
<tr>
<td>Rural (Outlying Town/Remote)</td>
<td>54.60</td>
<td>74.86</td>
</tr>
</tbody>
</table>

**Free and reduced meals rate of schools.** The next demographics question to be explained and compared to the Colorado demographics information is the Free and Reduced Meals (FARM) rate of the schools in which the principals work. In Table 4.5 below the researcher showed CDE and the participants percentages for comparison.

Table 4.5

*FARM rate: state of Colorado verses participants*

<table>
<thead>
<tr>
<th>Percent of students on FARM for Schools</th>
<th>Number of Participants</th>
<th>Percent of Participants</th>
<th>Colorado State Percentages (CDE, March, 2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9</td>
<td>8</td>
<td>5.30</td>
<td>9.87</td>
</tr>
<tr>
<td>10-19</td>
<td>11</td>
<td>7.20</td>
<td>11.78</td>
</tr>
<tr>
<td>20-29</td>
<td>12</td>
<td>7.90</td>
<td>10.09</td>
</tr>
<tr>
<td>30-39</td>
<td>13</td>
<td>8.60</td>
<td>10.77</td>
</tr>
<tr>
<td>40-49</td>
<td>20</td>
<td>13.20</td>
<td>12.23</td>
</tr>
<tr>
<td>50-59</td>
<td>27</td>
<td>17.80</td>
<td>11.16</td>
</tr>
<tr>
<td>60-69</td>
<td>19</td>
<td>12.50</td>
<td>9.65</td>
</tr>
<tr>
<td>70-79</td>
<td>20</td>
<td>13.20</td>
<td>8.75</td>
</tr>
<tr>
<td>80-89</td>
<td>18</td>
<td>11.80</td>
<td>10.66</td>
</tr>
<tr>
<td>90-100</td>
<td>4</td>
<td>2.60</td>
<td>5.04</td>
</tr>
</tbody>
</table>
Schools with a FARM rate between 40 percent and 89 percent took up 68 percent of the participants for this study. All categories of FARM rates were represented in the participants.

**Size of school.** The size of school the principal was in charge of is the next demographic to be reported. Results can be found in Table 4.6. CDE also had these demographics for all Colorado Schools. Schools that had between 101 and 700 students represented 84.8 percent of all of the participants to take the survey questionnaire.

Table 4.6

<table>
<thead>
<tr>
<th>Size of School</th>
<th>Number of Participants</th>
<th>Percent of Participants</th>
<th>State of Colorado Percentages (CDE, March 2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-100</td>
<td>6</td>
<td>3.90</td>
<td>10.60</td>
</tr>
<tr>
<td>101-300</td>
<td>49</td>
<td>32.20</td>
<td>23.14</td>
</tr>
<tr>
<td>301-500</td>
<td>61</td>
<td>40.10</td>
<td>30.62</td>
</tr>
<tr>
<td>501-700</td>
<td>19</td>
<td>12.50</td>
<td>19.21</td>
</tr>
<tr>
<td>701-900</td>
<td>7</td>
<td>4.60</td>
<td>5.73</td>
</tr>
<tr>
<td>901-1100</td>
<td>3</td>
<td>2.00</td>
<td>3.39</td>
</tr>
<tr>
<td>1101-1300</td>
<td>3</td>
<td>2.00</td>
<td>2.28</td>
</tr>
<tr>
<td>1301-1500</td>
<td>0</td>
<td>0</td>
<td>0.96</td>
</tr>
<tr>
<td>1501-1700</td>
<td>1</td>
<td>.70</td>
<td>0.96</td>
</tr>
<tr>
<td>1701-1900</td>
<td>1</td>
<td>.70</td>
<td>1.38</td>
</tr>
<tr>
<td>1901-2100</td>
<td>1</td>
<td>.70</td>
<td>0.47</td>
</tr>
<tr>
<td>1201+</td>
<td>1</td>
<td>.70</td>
<td>1.17</td>
</tr>
</tbody>
</table>

**Type of school.** The next demographic was the type of school. CDE also had information regarding this demographic. The participants gave answers that fit into pre-arranged categories offered by the researcher of this study or filled in an answer for “other”. All “other” answers were coded into the categories below with the different answers placed into parentheses.
Table 4.7

Type of School: State of Colorado Verses Participants

<table>
<thead>
<tr>
<th>Type of School</th>
<th>Number of Participants</th>
<th>Percent of Participants</th>
<th>State of Colorado Percentages (CDE, 1999-2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary School (K-5, K-6, Pre k-5, etc.)</td>
<td>81</td>
<td>53.29</td>
<td>50.08</td>
</tr>
<tr>
<td>Middle School (5-8, 6-8, 7-8, etc.)</td>
<td>25</td>
<td>16.45</td>
<td>14.04</td>
</tr>
<tr>
<td>High School (Regular, Alternative)</td>
<td>27</td>
<td>17.76</td>
<td>16.59</td>
</tr>
<tr>
<td>Elementary/ Middle School (k-8, pre k-8)</td>
<td>2</td>
<td>1.32</td>
<td>8.74</td>
</tr>
<tr>
<td>Elementary/Middle/ High School</td>
<td>11</td>
<td>7.24</td>
<td>4.61</td>
</tr>
<tr>
<td>Middle/ High School</td>
<td>4</td>
<td>2.63</td>
<td>14.04</td>
</tr>
<tr>
<td>Other (District, Home School)</td>
<td>2</td>
<td>1.32</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 4.8

Final Effectiveness Rating – Sample Verses State of Colorado

<table>
<thead>
<tr>
<th>Final Effectiveness Rating</th>
<th>Number of Participants</th>
<th>Percent of Participants</th>
<th>State of Colorado Percentages (CDE, 1999-2013)</th>
<th>State Percent (Minus Not Rated) (CDE, 1999-2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly Effective</td>
<td>56</td>
<td>36.84</td>
<td>29.57</td>
<td>33.39</td>
</tr>
<tr>
<td>Effective</td>
<td>94</td>
<td>61.84</td>
<td>53.64</td>
<td>60.57</td>
</tr>
<tr>
<td>Partially Effective</td>
<td>2</td>
<td>1.31</td>
<td>5.18</td>
<td>5.84</td>
</tr>
<tr>
<td>Ineffective</td>
<td>0</td>
<td>0</td>
<td>0.16</td>
<td>0.20</td>
</tr>
<tr>
<td>Not Rated</td>
<td>0</td>
<td>0</td>
<td>11.45</td>
<td>0</td>
</tr>
</tbody>
</table>

Final effectiveness rating. The final demographics question asked centered on the effectiveness rating the participant received for the 2016-2017 school year. For this question, there were two options for those who had not been evaluated using CMES for principals.\textsuperscript{34} Those who chose either of those answers were filtered out not only for this

\textsuperscript{34} These two options were: I was evaluated, but not using CMES. I was evaluated, but not as a principal.
question, but for the whole survey questionnaire because they have not been a part of the CMES for principals (as principals) for a full cycle of the evaluation. The Colorado Department of Education also has statistics based on principals’ effectiveness ratings. The researcher for this study removed the “Not Rated” category from the CDE data, because the researcher did not include a not rated category for the final sample of participants.

**Conclusion to Response Rate and Demographics**

Although a 50 percent response rate was expected, the response rate for this online survey questionnaire was significantly lower at 16.02 percent. The researcher is satisfied with the data that were collected, however, because the details outlined about how the survey questionnaire was distributed, the attempt to ensure transferability through an attempt at maximum variation. CDE data were provided for demographics that matched the demographics collected by the researcher of this study. Next is the data collected that answered the research questions.

**Answers to the Research Questions**

This part of Chapter IV will include four sections. The first section will be some simple descriptive statistics based on the question on the survey questionnaire that states: How often do you receive evaluative feedback for the Colorado Model Evaluation System? The remaining three sections will all be a mixed methods analysis based on the three research questions. The first two research analyses will follow this sequence: 1. start with quantitative analysis, 2. Next comes the qualitative analysis, 3. The final analysis will be a short comparison between the two sets of information. The analysis of the final research question will start with a quantitative analysis and end with a mixed
analysis where the researcher sorted the qualitative data according to favorable, neutral, or negative answers to the quantitative question. From all three research questions, the researcher will give a thorough mixed methods analysis of the perceptions of the CMES and feedback given that is aligned with the CMES by sample principals.

**Descriptive Statistics for: How Often Do You Receive Evaluative Feedback for the Colorado Model Evaluation System?**

The range of frequency of feedback goes from as often as once a week to never getting evaluative feedback, as reported in Table 4.9. When looking at the “Cumulative Percent” category of Table 4.9, 30 percent of the principals reported receiving evaluative feedback as often as every two months or less. Out of 152 survey respondents, the largest two responses were once a year with 45 responses (29.6%) and once a semester with 58 responses (38.2%).

Table 4.9

*How Often Do You Receive Evaluative Feedback for the Colorado Model Evaluation System*

<table>
<thead>
<tr>
<th>Frequency of Feedback</th>
<th>Frequency of Responses</th>
<th>Percent of Responses</th>
<th>Cumulative Percent of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>1</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>Once a week</td>
<td>2</td>
<td>1.30</td>
<td>2</td>
</tr>
<tr>
<td>2-3 times a month</td>
<td>6</td>
<td>3.90</td>
<td>5.90</td>
</tr>
<tr>
<td>Once a month</td>
<td>20</td>
<td>13.20</td>
<td>19.10</td>
</tr>
<tr>
<td>Once every 2 months</td>
<td>17</td>
<td>11.20</td>
<td>30.30</td>
</tr>
<tr>
<td>Once a semester</td>
<td>58</td>
<td>38.20</td>
<td>68.40</td>
</tr>
<tr>
<td>Once a Year</td>
<td>45</td>
<td>29.60</td>
<td>98</td>
</tr>
<tr>
<td>I have never received evaluative feedback in my current position.</td>
<td>3</td>
<td>2</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>152</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
Research Question #1: What Are the Principals’ Perceptions of the Colorado Model Evaluation System Evaluation Process?

To describe the statistics for each of the research questions, the following format was used. First the answers for the Likert Scale questions were detailed and evaluated. Then the process and themes for the qualitative data is described. Finally, a summary is given for the entire research question.

Quantitative analysis. For the first research question the descriptive statistics were first analyzed. Of the 152 total respondents, 149 gave answers to this question. Of the total sample, 49.4 percent have a positive or extremely positive impression of the CMES. The most frequent response was “Positive” with 72 of the 149 marking this as their answer. The percent of people who had neither a positive nor a negative impression of the CMES was 40.1. This leaves 13 principals who had a negative or extremely negative experience with the CMES (8.7 percent). The median and mode for this question is positive. Though positive was the largest category, just over 50 percent of this sample had a positive perception of the CMES.

35 For this question: 1 the most positive answer, 2 is positive, 3 is a neither positive nor negative answer, 4 is negative, and 5 is extremely negative.
Table 4.10

What is Your Overall Perception of the Colorado Model Evaluation System for Principals

<table>
<thead>
<tr>
<th>Perception of CMES</th>
<th>Frequency of Responses</th>
<th>Percent of Responses</th>
<th>Cumulative Percent of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely Positive</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Positive</td>
<td>72</td>
<td>48.30</td>
<td>50.30</td>
</tr>
<tr>
<td>Neither Positive Nor Negative</td>
<td>61</td>
<td>40.90</td>
<td>91.20</td>
</tr>
<tr>
<td>Negative</td>
<td>11</td>
<td>7.40</td>
<td>98.60</td>
</tr>
<tr>
<td>Extremely Negative</td>
<td>2</td>
<td>1.30</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>149</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

For this research question, the researcher used both analysis of variance (ANOVA) and chi-squared tests to determine whether or not categories of different demographics answered the closed-ended research questions differently. An ANOVA was run for FARM rate, size of school, principals’ total years, and principals’ years in their current position. For the FARM rate, the categories were collapsed in order to provide larger numbers per category.\(^{36}\) The ANOVA for FARM rate was not statistically significant. School size was also collapsed when doing this ANOVA.\(^{37}\) The ANOVA calculation showed there is not statistically significant differences. Principal’s total number of years and years in their current position also did not show statistically significant differences.

Using this research question, the researcher also did an analysis on whether there is a statistically significant difference of the perceptions of principals according to gender

\(^{36}\) For FARM rate: 0-29% = 1, 30-59% = 2, 60+% = 3.

\(^{37}\) For school size: 1-300 = 1, 301-500 = 2, 501+ = 3.
and comparing those participants from the Front Range of Colorado with the rest of the state of Colorado. Using the Chi Square test for gender, it was found that there is no statistical significance between the experience of males and females with regards to this question. The same was true for the principals on the Front Range verses the rest of the state.

An analysis was also done that compared the answers given to this research question to the frequency of feedback that was analyzed in the previous section. The results of this analysis can be found in Table 4.11. In this analysis, it was found that as the frequency of the feedback decreased, so did the percent of participants who had a positive view of the Colorado Model Evaluation System. As the frequency of feedback decreases, the percent of respondents who answer *Neither Positive Nor Negative* increases and the percent of respondents who answer *Positive* decreases.

---

38 A table with the numbers for each of these percentages can be found in Appendix H.
Table 4.11

*Frequency of Feedback Compared to the Answers to Research Question Number One*

<table>
<thead>
<tr>
<th>Frequency of Feedback</th>
<th>Percent Extremely Positive</th>
<th>Percent Positive</th>
<th>Percent Neither Positive nor Negative</th>
<th>Percent Negative</th>
<th>Percent Extremely Negative</th>
<th>Total Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>n=1</td>
</tr>
<tr>
<td>Once a week 2-3 times a month</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=2</td>
</tr>
<tr>
<td>Once a month</td>
<td>33.34</td>
<td>50</td>
<td>16.66</td>
<td>0</td>
<td>0</td>
<td>n=6</td>
</tr>
<tr>
<td>Once every two months</td>
<td>5.26</td>
<td>52.63</td>
<td>31.58</td>
<td>10.53</td>
<td>0</td>
<td>n=19</td>
</tr>
<tr>
<td>Once a semester</td>
<td>0</td>
<td>57.89</td>
<td>35.29</td>
<td>0</td>
<td>0</td>
<td>n=17</td>
</tr>
<tr>
<td>Once a year</td>
<td>0</td>
<td>54.39</td>
<td>40.35</td>
<td>3.51</td>
<td>0</td>
<td>n=57</td>
</tr>
<tr>
<td>Never</td>
<td>0</td>
<td>34.09</td>
<td>47.73</td>
<td>15.91</td>
<td>2.27</td>
<td>n=44</td>
</tr>
<tr>
<td>Total</td>
<td>2.00</td>
<td>48.30</td>
<td>40.90</td>
<td>7.20</td>
<td>1.30</td>
<td>n=149</td>
</tr>
</tbody>
</table>

Qualitative data analysis. The quantitative research question then had four follow-up questions, which were open ended. Two of the questions asked about what was going well with the CMES and two asked for areas of growth. The researcher followed the process outlined in Chapter IV and combined the codes into four prominent categories. These categories were each given definitions by the researcher to make the connections between the category, those codes that fit under those categories, and ultimately connecting the comments by the participants. One set of categories answered:

39 The columns are the participants’ perceptions of Question #1. The Rows are the Frequency of Feedback.
40 Tables showing the numbers for the percentages represented in Table 4.11, 4.13, and 4.15 can be found in Appendix H.
41 To explain the calculations for Table 4.11 - 5.26% of the 19 principals who responded that they received feedback once a month had an extremely positive impression of the CMES.
What has gone well for CMES and What has been most helpful? The other set of categories answered: Where can CMES improve and What has it been least helpful?

**What has gone well for the Colorado Model Evaluation System?** There were two positively focused questions that asked: What has gone well with CMES? (and) What has been helpful with CMES? The answers to these two questions were coded separately first, then combined, and categories were created after the codes for the questions were combined. After analyzing all the comments, and going through rounds of coding, five categories became apparent: expectations, growth, comprehensive, conversation, and negative comments.

**Expectations.** For this category, expectations refers to that which the person being evaluated expects, or that which is expected of this person, during the evaluation process. 75 comments made by the participants fit in this category. The participants distinguished between expectations with comments related to expectations of a principal’s job and expectations of the system as a whole.

Related to the principal’s job expectations, many of the participants in this sample believed that the CMES provided strong standards of leadership that guides their work. Many of the answers in this regard tied directly to the standards that are in place in CMES. One of the comments that exemplified the appreciation of the standards was, “The principal standards are accurate and exemplify our reality as a building principal.” (Principal 110)\(^{42}\). More of the comments tied these standards into an appreciation for the rubric, and the specific elements and performance indicators within the rubric. “I like

\(^{42}\) Participants were assigned numbers according to research question. These numbers were included to provide context to the variety of principals who commented.
self evaluating knowing in the fall which Standards and Elements I need to focus on to continually grow” (Principal 106). The appreciation of the parts of the rubric could further be exemplified by, “I like the explicit elements that I can self-assess myself on and focus towards making improvements. The process is kind of a road map on being highly effective” (Principal 97). The expectations set by the rubric provide a guide for principal’s work.

Appreciation was also expressed, by the participants, in the expectations of the system as a whole. Multiple elements of the system were mentioned throughout many of the participants’ comments including Professional Growth Plan (goal setting), the rubric, the meetings, etcetera. However, most of the comments for this theme have to do with the appreciation of consistent expectations for the system. “There is a required structure with timestamps and procedures” (Principal 56). Other participants appreciated the consistent expectations for all involved in the CMES. “It provides consistency from year to year and from administrator to administrator” (Principal 66). One participant valued the scope of the tool. “I like a standardized system for our district as well as the state” (Principal 139). Some participants welcomed that the job someone occupies does not matter either. It is a “similar process to teachers so sends the message we’re all in this together” (Principal 59). Overall, the participants who appreciated the expectations of the system, liked that the system was the same no matter who you are or where you work.

*Growth.* The researcher’s definition of growth is the attempt by someone to improve themselves and the allowance by the system for the participants to get better. Over the course of both questions there were 56 comments that fit into the category “growth”. When mentioning how the CMES has been helpful or positive for principals,
the participants’ comments divided growth into two manifestations: the systems focus on improvement and the goal setting that comes with this improvement.

The participants liked that this system is based on improvement. “I like that it is based on a continuous cycle of improvement” (Principal 116). Having this improvement tied to the principal’s job was the focus of many comments of these participants. Participants referenced either the skills that were laid out by CMES or named the parts of the evaluation (rubric, standards, elements, indicators) as that which helped them attempt to improve. It provided, “overarching categories that help guide my thinking and growth” (Principal 72). For others “it specifically identifies diverse areas that principals can grow in to increase effectiveness in their school” (Principal 20). This instrument provided the opportunity for growth and specificity in where to improve.

The system also provided principals the opportunity to set targets for growth through goal setting. Many in the sample wrote specifically about writing goals or referenced the Professional Growth Plan which is based on goal setting. The system “Helps guide my goals” (Principal 124). Or, the system makes it so principals are “able to set goals and clear targets” (Principal 13). Even when goal setting was not specifically stated, CMES, “does give me a good guide for focusing on my own specific areas of improvement” (Principal 3). The system helps principals target their areas of growth.

Comprehensive. For this category, comprehensive means complete and thorough and it details and provides a guideline to all aspects of a principal’s job. Twenty eight comments fit into the category comprehensive and all of comments occurred in response to question 15, which stated: What do you like about CMES? Comments supported the rubric, standards, and indicators and that it had some details of a principal’s job,
including how to be exemplary. Most of the comments were similar to the following three. Referencing the overarching expectations of principals; “I like that it takes into consideration different standards of leadership. It does give me a good guide for focusing on my own specific areas of improvement” (Principal 3). In regard to the specificity and details of the rubric; “I like that it gives you indicators that quantify our work” (Principal 5). One participant made a comment regarding having clear and comprehensive evaluation system: “I like the rubric and the performance being very clear about what I am evaluated for” (Principal 138). Overall the participants whose responses fit into this category appreciated having a comprehensive system that details their jobs.

Conversation. For this category, conversation means that the principals have had conversations with others that are formative to their practice. Comments coded into this category numbered 26 times over the 2 questions. The conversations mentioned by participants took on two forms; namely, conversations principals are having with their evaluators and those they evaluate and CMES providing a conversation focus.

Many of the participants who mentioned the importance of conversations discussed their conversations with others. Most of the time the conversation with the evaluator was mentioned sometimes by position and sometimes just by using the word evaluator. Comments like “being able to discuss professional areas with my evaluator” (Principal 37); “It encourages communication and collaboration with my evaluator” (Principal 65); and “It structures a conversation” (Principal 56) illustrate the idea that the participants appreciate the opportunity to engage in conversations with their supervisors about their performance.
Other participants remarked on having the conversation as a structure and what this structure provides. General words like structure, talking points, and framework were used in reference to CMES’s structure of the communication. “It has great talking points and descriptors for reflective conversations” (Principal 67). “It creates a framework to talk through evaluations” (Principal 90) points to how the system is providing a structure to have conversations.

Specific parts of the evaluation system were mentioned because of how they facilitate a conversation. Structure of the rubric, Mid-Year Review (MYR) and End of Year Review (EYR), and Standards were all explicitly mentioned. Regarding the structure the rubric provides: “I appreciate the conversations around growth facilitated by the rubric” (Principal 48). Having two points in the year where there is a required conversation: “Writing goals and having mid-year and end of year conversations about them” (Principal 63). Regarding standards: “They [conversations] align well with our target standards” (Principal 90).

**Negative comments.** Although these two questions both asked for positive comments, there still were those who were frustrated with the system and did not provide positive responses. There were 13 responses that were considered negative responses to questions asking for positives over two questions. Six of the people who had negative responses either filled negative comments for both positive questions, or did not answer the other positive question. Most of the negative responses were simple words like “Nothing” or “None” (Principal 28) when asked about positives. Their feedback on the instrument was all about how it could improve (next section). One comment was about how old (locally created) evaluations were better, “Locally developed evaluations of the
past were less cumbersome, more efficient and more helpful. CMES is not helpful in any way” (Principal 42). One complimented the superintendent, “The conversation with the Superintendent is helpful, not the form itself” (Principal 124). One participant mentioned that the system was not helpful and, “it would be more helpful to me if the language was clarified and redundancies removed” (Principal 40). One participant wrote that the amount of time needed for the CMES impacted them negatively “With so many other plans, supports, resources and initiatives it tends to take a back seat until the end of the year” (Principal 131). This group of negative comments is a minority with 13 total comments.

Mixed analysis for negative comments. A further analysis was done on the negative comments. Nine participants were found to have given all the negative comments. For all nine of these participants, the researcher looked at how they answered: What as your Final Effectiveness Rating for the 2016-2017 School Year? & How often do you receive evaluative feedback for the Colorado Model Evaluation System? The answer for the Final Effectiveness Rating was: three participants were highly effective, five of the participants were effective, and one of the participants was partially effective. These numbers were comparable to the overall sample. For the question about frequency of feedback: one participant answered – once a month, three participants answered – once a semester, and five of the participants answered – once a year. In short, the negative perceptions of the CMES process articulated in these findings were not directly correlated with the individual rating of a participant nor to the amount of feedback the participant received, when comparing this sample of nine participants to all participants.
**Where or how can the Colorado Model Evaluation System improve?** For the two questions that asked about the concerns participants had with the CMES, there were five predominant categories. Four categories had the highest number of comments associated with them: clarify and streamline the system, my job, school district use of the system, and Measures of Student Learning. The final category has to do with answers given that point to ignoring the system or making the system disappear. These categories provide a comprehensive understanding of the perspective of the participants with where they had concerns with the system.

**Clarify or streamline the system.** The definition of this category is that there are parts of the system that need to be cleaned up to improve the overall quality of interaction the principals have with their evaluators and the rubric as a part of the CMES system.

One hundred and eleven comments were coded and then placed into this category. Comments coded under this category showed that participants sought clarification on the system and the rubric.

The streamlining of the system mostly was explained by the amount of time and extra work that principals are required to put in to follow the expected steps. A specific example of this came from Principal 3, “Too time intensive. Less of the formality with, ‘I submit’… Now you submit now you accept, etc.” Principal 42 stated that “This system is cumbersome and lacks efficiency.” Many participants also put length as the issue without sufficient clarification on what was the issue.

The focus of most of the answers that followed this category came with critical comments about the standards, rubric, elements, and indicators. First, the suggestion of many was that the indicators need to be narrowed down. Some of these comments were
made in comparison to what has happened with a pilot being done for the teacher rubric.43 “Make a similar effort as was made on the teacher rubric to make it shorter, more practical to use” (Principal 6). Other participants wrote how the number of indicators and elements impacted the quality of the evaluation. “There are some redundancies and too many attributes to consider. It would be more impactful if it were more focused on fewer indicators” (Principal 35). “Narrow down the indicators” (Principal 5). “It can be overwhelming because of so many elements” (Principal 7). Other participants believed the indicators provide too much subjectivity, are not clear, or are confusing, and that clearing these problems up can improve the rubric. These different points can be summarized by the following:

The language of the standards and elements can be somewhat confusing and ultimately the person doing the evaluation still has subjectivity – Do I do this 100% of the time or 50% of the time? Have I done it once or do I do it every day? (Principal 63).

This statement shows an ambiguity of wording and of frequency.

My job. The definition of the second category is that the focus of the evaluation needs to be specific to the job of the person being evaluated. Comments related to this category numbered 41 between both questions. Specifically, the answers related to my job centered on parts of the system that people might have trouble connecting to their jobs and the lack of definition for artifacts.

---

43 Explanation for the pilot of the revised teacher rubric can be found at: http://www.cde.state.co.us/educatoreffectiveness/statemodelevationsystem-pilot
Participants expressed concern that there are elements within the rubric that did not represent their specific job or were impossible to attain. A few participants mentioned that the system did not relate to their jobs at all. “Evaluation process. [It is] Not applicable to my normal mode of operation” (Principal 5). Others felt like there were specific parts of the evaluation that do not fit their specific job or their day to day operations. “Some benchmarks are difficult if not impossible to obtain – for example PK-20 if we just work at one level” (Principal 45). Others had concerns about expectations with in the rubric like, “as a principal with no assistant principal at a title one building, I do not often have extra time to sit on district committees, etc. and I feel these things should not be required to be an accomplished principal” (Principal 68). Some of the participants mentioned assistant principals specifically as being evaluated using a rubric for assistant principals. These participants felt there should be a, “Different rubric for Aps and principals. Their work is very different and principal doesn’t capture all that an AP should and does do” (Principal 59). Whether the evaluation tool did not connect with participants’ jobs in general, had expectations that felt unreasonable for their position, or spoke to the challenges of having the same rubric for two different positions (assistants verses principals), there were concerns whether the rubric set reasonable expectations for the jobs being performed.

Participants were also confused by what could be included as an artifact. A few wanted resources demonstrating exemplars for the elements and indicators. One example of this was, “a resource guide similar to what CDE provides teachers” (Principal 96). Some suggestions were not along the lines of a whole resource guide, but at least “more
explanation on examples of what elements mean” (Principal 136). These principals wanted examples to help better understand the expectations stated in the rubric.

*School district use of tool.* The definition of this category is the evaluator’s interaction with the evaluation system. This category was mentioned 16 times between the two questions. The first concern raised by the participants in the sample had to do with the poor job principals’ evaluators are doing with the CMES as a whole. “If my evaluator does not use the system well, then the use of the system is not meaningful” (Principal 9). More principals were concerned about the feedback they were receiving, with most who responded with this theme saying they are not receiving enough feedback. One stated that there is a “lack of on-going conversations to improve and celebrate performance related to the CMES” (Principal 99). Three of the participants wanted district expectations on use of the system. Two of the participants wanted expectations based on use of the system, and one wanted the evaluators to have more training “to create clarity and consistency” (Principal 151).

*Measures of student learning.* The Measures of Student Learning are the data collected that determines student growth or achievement. In Colorado, these data are supposed to be at least 50 percent of the Final Effectiveness Rating Educator Effectiveness Act (State Council for Educator Effectiveness, 2011; Educator Effectiveness Act). This category was stated in some capacity 14 times in the answers to the questions asking for areas that this sample is struggling with while participating in the CMES. Most of the principals whose responses fit in this category just named MSLs as being the least helpful part of the evaluation. Two participants both had concerns about the impact MSLs have on Final Effectiveness Ratings. One was concerned about, “Being
penalized because my school is at high risk feels unhelpful” (Principal 20). Another participant stated, “the MSL growth piece is a joke and needs to go away. It actually is minimizing the importance of professional practices and is inflating evaluations” (Principal 11).

Disappear and non-priority. A small minority saw no redeeming value in the CMES. Four participants answered in this way. They either wrote that they wanted the system to stop existing or that it was of no priority to them that they just ignored it and continued their work. One specifically suggested to “Get rid of it. It has not had any impact on how or what I do professionally. It is nothing more than another hoop to jump through” (Principal 28). For a small minority, after being asked for suggestions of improvement, they could only suggest that it go away.

Conclusion to research question number 1. The participants stated that they had a generally positive impression of the CMES system. Those who had neither a positive nor a negative impression of CMES were 40.1 percent of the respondents. So, although the median and mode for this question came to positive, there were a large number of participants who answered the question with a neutral response. In addition, there was an increase in neutral perceptions of the CMES when the frequency of feedback received by the participant decreased.

With an understanding that the overall participants’ impressions of the CMES were positive, there were five themes derived from answers to the questions which asked what the participants liked about CMES that became apparent: expectations, growth, comprehensive, conversation, and negative responses. When further breaking down the negative responses, it was found that there was a higher percentage of this group of
participants who received feedback once a year compared to the overall sample, suggesting that more frequent feedback might alter participants’ perceptions. There were also challenges within the CMES. The themes that came from these challenges were: clarify and streamline the system, my job, district use of tool, Measures of Student Learning, and disappear. An examination of the second research question followed a similar format to the first. The explanation of the results for the second research question follows.

**Research Question # 2: What Are the Principals’ Perceptions of the Colorado Model Evaluation System Evaluation Feedback?**

Much like the first research question, this research question had two parts to its analysis followed by a conclusion that combines both results. The analysis starts with quantitative analysis to show the results of the closed ended question with Likert answers. This is followed by the qualitative analysis which shows the results of the coding of the two follow up open-ended questions: How has the feedback you have received as a result of the CMES been helpful? and How could the feedback as the result of CMES be improved? These themes were given general definitions to clarify how they became themes. In the conclusion is a short synopsis of both results comparing both qualitative and quantitative data.

**Quantitative analysis.** For the second research question, the descriptive statistics were first analyzed. Of the 152 total respondents, 144 gave answers to this question. Of the total sample, 56.3 percent have a positive or extremely positive impression of the

---

44 For this question: 1 the most positive answer, 3 is a neither positive nor negative answer, and 5 is the most negative answer.
feedback they are receiving because of CMES. The most frequent response was “Positive” with 50.69 percent of sample principals marking this as their answer. The percent of principals who had neither a positive nor a negative impression of the CMES was 41.0. Only 2.8 percent of the principals had a negative or extremely negative experience with the feedback they are receiving according to CMES. The median for answers to this question was that the participants had a “positive” impression of the feedback they were receiving using the CMES. Over 50 percent of the participants responded positively with the second largest percent of answers being the neutral answer.

Table 4.12
What Has Been Your Perception of the Feedback You Have Received as a Part of the Colorado Model Evaluation System

<table>
<thead>
<tr>
<th>Perception of the Feedback Using the CMES</th>
<th>Frequency of Responses</th>
<th>Percent of Responses</th>
<th>Cumulative Percent of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extremely Positive</td>
<td>8</td>
<td>5.60</td>
<td>5.60</td>
</tr>
<tr>
<td>Positive</td>
<td>73</td>
<td>50.70</td>
<td>56.30</td>
</tr>
<tr>
<td>Neither Positive Nor Negative</td>
<td>59</td>
<td>41</td>
<td>97.20</td>
</tr>
<tr>
<td>Negative</td>
<td>3</td>
<td>2.10</td>
<td>99.30</td>
</tr>
<tr>
<td>Extremely Negative</td>
<td>1</td>
<td>0.70</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Just like research question one, a univariate ANOVA and the chi-squared test were used to see if any difference existed between experiences from different demographics. For the demographics where ANOVA was used, FARM rate, principal total years, and principal in current position all were not statistically significant. The school size demographics was approaching significance at .06. When the researcher looked further into this data, the difference between the schools with 1-300 students was
not significantly different than those with 301-500 students, but was significantly different than schools with 501+ students.

A bivariate analysis on whether there is a statistically significant difference of the perceptions of the participants according to gender and the experience of Front Range principals versus the rest of the state of Colorado was run. Again, after using the Pearson’s Chi Square test for gender, it was found that there is no statistical significance between the experience of males and females with regards to this question. The same was true with the location of the school district (Front Range versus the rest of the state of Colorado).

A comparison was done between how participants answered this research question and the frequency of feedback they reported receiving as a part of the CMES process. The only category that seemed to vary both from the rest of the categories and the answers given by the whole sample to this question were those participants who reported receiving feedback as a part of the CMES once a year. The numbers are summarized in Table 4.13.
Table 4.13

<table>
<thead>
<tr>
<th>Frequency of Feedback</th>
<th>Percent Extremely Positive</th>
<th>Percent Positive</th>
<th>Percent Neither Positive nor Negative</th>
<th>Percent Negative</th>
<th>Percent Extremely Negative</th>
<th>Total Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=0</td>
</tr>
<tr>
<td>Once a week</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=2</td>
</tr>
<tr>
<td>2-3 times a month</td>
<td>16.66</td>
<td>83</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=6</td>
</tr>
<tr>
<td>Once a month</td>
<td>10.00</td>
<td>50.00</td>
<td>30.00</td>
<td>10.00</td>
<td>0</td>
<td>n=20</td>
</tr>
<tr>
<td>Once every two months</td>
<td>5.88</td>
<td>52.94</td>
<td>41.18</td>
<td>0</td>
<td>0</td>
<td>n=17</td>
</tr>
<tr>
<td>Once a semester</td>
<td>3.77</td>
<td>62.26</td>
<td>33.96</td>
<td>0</td>
<td>0</td>
<td>n=53</td>
</tr>
<tr>
<td>Once a year</td>
<td>4.65</td>
<td>32.56</td>
<td>58.14</td>
<td>2.33</td>
<td>2.33</td>
<td>n=43</td>
</tr>
<tr>
<td>Never</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>n=3</td>
</tr>
<tr>
<td>Total</td>
<td>5.30</td>
<td>48.00</td>
<td>38.80</td>
<td>2.00</td>
<td>.70</td>
<td>n=144</td>
</tr>
</tbody>
</table>

Qualitative data analysis. The quantitative question was then followed up with two open ended qualitative questions. The researcher posed one question that asked how the feedback has been helpful. The second question asked for how feedback could be improved. Two sets of categories were determined based on the answers from the two different questions.

How has the feedback using Colorado Model Evaluation System been helpful?

Four categories were found from the answers to this question: improvement, school district evaluator specific suggestions, positive feedback, and non-productive or no

---

45 The columns for Table 4.13 are participants answers to Research Question #2. The Rows are the participants answers to the frequency of feedback they are receiving from their evaluator.
feedback. These categories are placed in order according to the number of comments that fit under each category, with the largest coming first. The category with the most mentions by participants was improvement.

*Improvement.* Improvement was defined by the researcher as the feedback according to CMES being supportive of a system of improvement. There were 50 participants who mentioned improvement or talked about how the feedback helped to provide focus in order to improve.

Many participants stated that the intent of improving practice was an important point. Words like “next steps,” (Principal 27) and “growth” (Principal 48) were used and really could be exchanged for the word improvement throughout these comments. “It provides me with next steps in my professional growth” (Principal 27), hints that the feedback as a part of the evaluation system provides structure within which to grow. Some principals appreciated that the feedback facilitates growth. “It’s allowed me to learn and grow as a professional” (Principal 102). Most of the comments by principals under this topic, however, mentioned the word improve or improvement along with parts or the whole system of feedback as that which has been most helpful. One such comment was, “There are specific behaviors or targets I could improve upon according to CMES that help me from getting too complacent” (Principal 143).

Improvement, according to the participants, also focused on the feedback provided to enable the principals to grow in their profession. For some of the sample participants the focus was occurring through goal or target setting. Many named “set

---

46 The exception to this was that the negative feedback, with 22 times mentioned, was placed last.
goals” (Principal 23) specifically or used a version of the word “target” (Principal 37) in place of the word goal. Many of the comments aligned with the following comment: “[The feedback] Allowed me to focus on clear goals and outcomes” (Principal 65).

Though most probably referencing goals or targets, some wrote comments similar to, “[Feedback] has given me a specific area to focus on” (Principal 5). Two participants also referenced specific indicators within the rubric that they were trying to improve on. One stated, “It’s made me realize that family and community engagement needs to be an area of focus while instructional leadership is an area of strength” (Principal 59).

Participants appreciation of a focus on growth was a strong sub-category for improvement.

School district/evaluator specific suggestions. The best way to define this category is that some of the participants connected the feedback specifically to their principal evaluator. There were 12 participants whose comments fit under this specific category. Whether through specific stories or through simple answers, the participants who answered using this category really felt their experience depended on the evaluator. A couple comments explicitly spoke of the system of feedback being dependent on the evaluator. “whatever tool is used for evaluation will be adjusted to meet the needs and expectations of the evaluator” (Principal 11). A few others wrote of their positive or negative experience that they have had with the feedback that was less CMES dependent and more evaluator dependent. One participant stated:

The prior superintendents provided quite positive feedback and help during my evaluations. They gave me effective and highly effective ratings, but provided ways to improve. The [most recent] superintendent did not provide feedback and
asked me to evaluate myself. I evaluated my performance as true to the rubric as I could and rated my performance as “Partially Effective.” The superintendent said if that is what I wanted then he’d go with the rating. I did not receive any feedback on if he agrees with my rating and no feedback on how to improve. I was told this year that whatever I decided my ratings would be were the ones I would receive. (Principal 154)

This experience encompasses the theme district/evaluator specific suggestions, because the feedback is evaluator dependent.

*Positive feedback.* This category is best described by the feedback being positive, and principals being evaluated appreciating that the feedback is positive. This category had 10 responses that best represent the above-stated definition. Participants described positive feedback from two perspectives: the evaluator is positive and the feedback reinforces the good that the principal is doing.

Those participants whose answers fit under this category, identified the evaluator being positive as significant and they were appreciative of the positive demeanor of the evaluator. In a small number it did not matter what the positive comments related to, they just appreciated the feedback being positive. This can be represented by the following comment: “My evaluator is very positive and knows I will do anything to help staff and students. She isn’t stuck on the many outcomes on the evaluation” (Principal 39). Positive evaluators and positive feedback was appreciated by those who gave these answers.

Related to the feedback and how it reinforces the principals and their work, one participant stated, “Just to know that I am making the right decisions” (Principal 128).
Another participant detailed experiencing positive feedback as, “The feedback from my evaluator has been positive and offered suggestions to improve my practice” (Principal 76). These statements encompass this theme in the best manner.

Non-productive, not helpful, or no feedback. This final category was created from negative answers given by the participants. The category can be defined as the feedback according to these participants either was not received or was counterproductive. Twenty-six of the participants made comments that fit under this category.

A few of the participants mentioned that they did not receive much (or any) feedback through the CMES. Some participants were explicit that the lack of frequency made the feedback not helpful. “I don’t receive much feedback so it is difficult to answer this question” (Principal 9). Or, “It was not of significant frequency to feel as helpful as it could have been” (Principal 20). Both of these participants wrote of the need for more frequent feedback to improve practice.

The participants also felt that feedback might have been received, but was not helpful using CMES. This came across in a negative way and a positive way. One principal stated, “The very little feedback I have had has not been helpful at all” (Principal 28). Two comments connected the positive to the evaluator and not to the CMES instrument. One stated, “The superintendent is very skilled in coaching and supporting staff, not necessarily due to CMES” (Principal 124). Both the positive and negative show that this group of principals feel that the feedback from the CMES system has not been helpful.

Non-productive, not helpful, or no feedback: a mixed analysis. For this category, a mixed analysis was done in order to find how the participants who answered with
negative comments answered the questions: *What as your Final Effectiveness Rating for the 2016-2017 School Year?* and *How often do you receive evaluative feedback for the Colorado Model Evaluation System?* The Final Effectiveness Ratings for the 152 of the sample verses the 26 who answered with negative comments for this question were similar. There was a difference between the 152 participants and these 26 participants when it came to the amount of feedback. Those who received feedback once a year represented 29.6 percent of the total participants, but 53.84 percent of the participants with negative answers. Those who reported receiving feedback once a semester represented 38.2 percent of the participants and 42.31 percent of the participants with negative answers. Those who reported negative comments for this question reported receiving a lower frequency of feedback than the sample of 152 participants.

*How can the feedback as a result of the Colorado Model Evaluation System be improved?* When analyzing the answers given for this question, four themes became apparent: specificity of feedback, frequency and timeliness of feedback, instrument or system based suggestions, and evaluator dependent. The themes are organized according to numbers of participants who had answers about this theme from largest to smallest theme.

*Specificity of feedback.* This theme can be defined in this context as whether or not the feedback that principals are receiving really connects with or impacts their work as a principal. This theme was mentioned 27 times. Participants’ viewed feedback from the perspective of feedback in general, and the need for specific feedback to relate to the rubric and their job.
Nine participants used the word specific in particular. A couple of the participants only stated, “more specific” (Principal 121). The majority of the sample who answered asking for specificity; however, asked for a better connection to either their job in general, their job duties specifically, and/or the rubric most specifically. Some participants wanted feedback that related to their specific job or school. Principal 5 explained this as, “Actionable feedback does not always relate to my work.” Others with general job concerns were wanting feedback that better related to their type of principal. “It would be nice if it matched what we do in our small, rural district” (Principal 33). Others broke their needs down further by connecting to their job duties. “Specific examples of quality action steps to take in order to grow as a leader in my building” (Principal 67). The final group of principals wanted specific feedback on how to improve based on the rubric. “Quantifying some of the rubric elements” (Principal 144). Overall, specificity of feedback on how to improve principals’ practice was a strong theme.

*Frequency and timeliness of feedback.* The second most frequently coded category for how to make the feedback more helpful can be defined as the amount of times that feedback is actually given to the principals. Comments that fit under this category were mentioned, in some capacity, 21 times.

Those participants who wrote about the frequency of feedback mostly spoke to not receiving enough feedback. Most comments could be explained by “frequent, meaningful feedback with the evaluator” (Principal 40). A couple participants wanted either less feedback or less time spent on the process. “The time consuming portion makes it harder to have more frequent conversations” (Principal 69). Overall those who answered under this theme wanted more and higher quality feedback.
A couple participants stated that more timely feedback would be helpful. One stated that “it could be more timely” (Principal 48). Another participant further explained their frustration with “We preach timely feedback for students but do we practice that with adults?” (Principal 78). These participants needed the feedback to come closer to when the event for which feedback is given occurs.

Instrument or system based suggestions. The following category can be defined as participants giving specific suggestions for the instrument itself or the setup of the system. Seventeen comments were coded under this category. This category is differentiated into changes suggested for the system as a whole and rubric specific suggested changes.

Many of the participants who made comments coded under this theme made suggestions to changing the CMES feedback system as a whole. Responses varied from how conversations are setup and handled to if the whole system is manageable. “The state could make the process more manageable for building leaders. We know the greatest resource is human capital so interaction is key” (Principal 89). Or, “Make it less of a checklist” (Principal 74). Some suggestions would be simple fixes. “In the goal setting section, it only addresses the barriers that I face and how I am going to address them. It does not ask to identify the positive things that I have done while I am working towards my goals” (Principal 63). According to the participants who answered with this category, systemic improvements need to be made.

Participants also commented on the instrument or rubric and offered suggestions. The comments varied from the impossibility of reaching standards to narrowing the elements and indicators. The statement “Narrow the elements” from Principal 87 gives a
general suggestion to shortening the rubric. “Most of the targets I didn’t hit are things that I knew I wouldn’t hit because they are unrealistic. Terms like ‘every student’ are unrealistic in my mind” (Principal 101). One participant’s suggestion encompassed most of the others’ comments. “we do not discuss the quality of the elements as often as I would like” (Principal 107). One participant made a suggestion that was a sub-category in an earlier question. “If the feedback was differentiated for the positions of Principal and Assistant Principal, it would benefit all evaluated” (Principal 97). According to these participants, rubric improvement would ensure higher quality conversations and potentially an overall improved experience for principals.

Evaluator dependent. The final theme for this question ties directly into the evaluators actions while providing feedback for the CMES. Although many of the comments and themes have to do with the evaluator’s interaction with CMES and the principals as those being evaluated, this theme is separate because evaluators were mentioned explicitly. This theme was mentioned 11 times.

Many of the participants wished for the evaluators of principals to adhere to expectations or have a certain level of understanding of the system. According to those who answered in this way, evaluators need to meet the expectations of providing effective feedback. These thoughts were best summarized, “I believe that CMES is a good system as long as those using the system are will[ing] to use it [as] an effective tool, otherwise it is just an exercise in clicking the right buttons by the specified deadline” (Principal 109). Another wrote about accountability of evaluators and training of evaluators. “Our evaluators need to be held to the same expectations that principals are with the teacher rubric” (Principal 9). One participant hinted at the lack of connection to
the principal or training of the evaluator. “Evaluators need to be more familiar with the role of the person they are evaluating” (Principal 64). According to these participants, evaluators need to be better connected to the evaluation system whether that is by engaging more with the system or getting a better understanding of the system.

Participants also stated that the overall effectiveness of the evaluation depends on the evaluator. “I think this is based on the evaluator and how they use the system” (Principal 37). Another talked about how inconsistency can happen in district and statewide. “I think it is dependent on each evaluator. It is tough to get consistency from each school in the district let alone statewide” (Principal 4). The effectiveness of the CMES evaluation feedback depends on the evaluator.

**Conclusion to research question number 2.** Much like the first research question, the impression of the feedback that this sample was receiving was positive. The median and mode of these Likert Scale answers was that the participants have a positive impression of the feedback they received using the CMES. This question also had a significant percent of the sample who answered “Neither Positive Nor Negative” with 41 percent. Though the Likert Scale answers skewed positive on this question there still was a significant number who decided to answer neutral. When comparing the results of this question to the frequency of feedback reported in an earlier question in the survey questionnaire the largest difference between the total sample population and the participants broken down according to frequency of feedback came from those who reported getting feedback one time a year.

With the majority of the answers given by the sample being positive there were four strong themes that came from the question requesting positive answers. These
themes are: improvement, district/evaluator specific themes, positive feedback, and non-productive, not helpful, or no feedback. The non-productive theme was further analyzed. From this analysis, it was found that the largest difference between the categories of feedback frequency and the total sample was for those who reported receiving feedback once a year. For the ways to improve the feedback, the sample gave answers that fell into four main themes: specificity of feedback, frequency and timeliness of feedback, instrument and system-based suggestions, and evaluator dependent.

**Research Question #3: According to Principals’ Perceptions, How Is the Feedback from the Colorado Model Evaluation System Evaluation Process Altering Principals’ Practice?**

The third and final research question to be analyzed is: According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principal’s practice? The original quantitative analysis will be the same as the two previous research questions. The qualitative/mixed analysis will follow what Creswell (2012) described as a “convergent design analysis – consolidating results” (p.551). The qualitative data were broken into three different categories of data according to the answers given by the participants for the quantitative question: positive impact, no impact, or negative impact. The reason behind breaking up these three categories is because each of these groups have had different experiences according to the answer given to the quantitative question, and the three sets of data will have different themes. The qualitative data was then coded and broken into categories according to answers to the quantitative question.

**Quantitative data analysis.** As was the case for the previous two research questions, descriptive statistics were first analyzed. One hundred and forty of the 152
respondents gave answers to this question. Those in the sample that had the most positive change in their practice numbered 60.8 percent. The percent of participants who have not had their practice impacted was 37.1 percent. Those whose practice had been negatively impacted take up 2.1 percent of the sample population. “Positive” was, again, the most frequent answer given with 57.9 percent of the participants marking this as their answer. The median for this question was also the answer “positive”. Much like the first two quantitative research questions, the answer to this question tended to be positive.

Table 4.14

<table>
<thead>
<tr>
<th>Feedback using CMES changed practice.</th>
<th>Frequency of Responses</th>
<th>Percent of Responses</th>
<th>Cumulative Percent of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greatly Changed (Positive)</td>
<td>4</td>
<td>2.90</td>
<td>2.90</td>
</tr>
<tr>
<td>Changed (Positive)</td>
<td>81</td>
<td>57.90</td>
<td>60.70</td>
</tr>
<tr>
<td>Has not changed my work.</td>
<td>52</td>
<td>37.10</td>
<td>97.90</td>
</tr>
<tr>
<td>Changed (Negative)</td>
<td>1</td>
<td>0.70</td>
<td>98.60</td>
</tr>
<tr>
<td>Greatly Changed (Negative)</td>
<td>2</td>
<td>1.40</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>140</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As was the case in the previous questions, ANOVA and chi-squared tests were run for some demographics to determine if any statistically significant differences occurred. The calculated ANOVA for FARM rate, school size, total years as a principal, and years in current position all were insignificant. The same was true of both the

47 For this question: 1 the most positive impact, 2 positive impact, 3 is no impact, 4 is negative impact, and 5 is extremely negative impact.
demographics where chi-squared tests were used (location = Front Range verses rest of the state of Colorado & gender of the principal).

A comparison was made, again, between the frequency of feedback and the Likert answers to this research question. The results can be found in Table 4.15. The largest discrepancy of percentages between all participants’ answers and the percentages according to each category of frequency of feedback came for those in the sample who answered that they were receiving feedback according to the CMES once a year. This is similar to the result from the previous question.

Table 4.15

<table>
<thead>
<tr>
<th>Frequency of Feedback</th>
<th>Percent Greatly Changed (Positive)</th>
<th>Percent Changed (Positive)</th>
<th>Percent Has not changed my work</th>
<th>Percent Greatly Changed (Negative)</th>
<th>Percent Changed (Negative)</th>
<th>Number of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>n=1</td>
</tr>
<tr>
<td>Once a week</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=2</td>
</tr>
<tr>
<td>2-3 times a month</td>
<td>33.34</td>
<td>66.66</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=6</td>
</tr>
<tr>
<td>Once a month</td>
<td>5.26</td>
<td>57.89</td>
<td>31.58</td>
<td>21.05</td>
<td>0</td>
<td>n=19</td>
</tr>
<tr>
<td>Once every two months</td>
<td>0</td>
<td>62.50</td>
<td>37.50</td>
<td>0</td>
<td>0</td>
<td>n=16</td>
</tr>
<tr>
<td>Once a semester</td>
<td>1.92</td>
<td>65.38</td>
<td>28.85</td>
<td>0</td>
<td>3.85</td>
<td>n=52</td>
</tr>
<tr>
<td>Once a year</td>
<td>0</td>
<td>46.34</td>
<td>53.66</td>
<td>0</td>
<td>0</td>
<td>n=41</td>
</tr>
<tr>
<td>Never</td>
<td>0</td>
<td>33.34</td>
<td>66.66</td>
<td>0</td>
<td>0</td>
<td>n=3</td>
</tr>
<tr>
<td>Total</td>
<td>2.90</td>
<td>57.90</td>
<td>37.10</td>
<td>0.70</td>
<td>1.40</td>
<td>n=140</td>
</tr>
</tbody>
</table>

Qualitative/ mixed methods data analysis. As was stated earlier in the chapter, the data for this final research question have been divided according to type of answers:

---

48 The columns represent the participants’ answers to Research Question #3. The
Changed/ Greatly Changed (Positive), Has not changed my practice, and Changed/ Greatly Changed (Negative). The order of the following data is according to frequency of answers. Positive answers had the largest number. No change in practice was the second largest category. Negative change to practice was the final set of data to be analyzed. The final categories are detailed below.

**Feedback from the Colorado Model Evaluation System has had a positive impact on practice.** Those participants with a positive change in their practice due to the CMES numbered 86. From this sub-sample of 86 participants, there were four categories that came from their answers to the open-ended question. These four categories are: focus, improvement, evaluator dependent, and reflection. There were many answers provided where both focus as a category, improvement as a category, and/or reflection as a theme crossed over, but these three themes had enough differences that they needed to be separated.

**Focus.** For this category, focus means that the feedback through the CMES helps to provide direction towards which principals should work. Forty-seven of the participants who stated that feedback had a positive impact on their practice had comments coded under this category. According to the participants, focus includes goal setting and specific areas of focus.

Related to focus and goal setting, participants mentioned goal setting and the professional growth plan explicitly and when the sample principals wrote of path and guidelines that were set by this process. Many of the participants wrote of developing

---

49 The numbers given to the principals will be different than the previous answers. Ex. These principals will be assigned numbers 1-86 randomly after they were sorted according to their positive answers to the closed question with Likert Scale answers.
targets to work on for their own growth. One specific example was, “My supervisor and I were able to converse about what is going well and what steps I can take to improve targeted areas. We are working together to help me meet my professional goals which has been helpful” (Principal 39). One participant generally appreciated the process. “It gives me something to reflect on and ideas of what my goals should be” (Principal 10). Participants also gave specific parts that they appreciated in conjunction with the goal setting. “Through my own self-assessment, reflection and goal setting, I have used the CMES to ensure that the work of myself and team is tied to student performance and teacher growth; along with working with our school community” (Principal 7). Goals and goal setting helped these participants to focus on their work.

The word focus was a word used by many participants to explain why the feedback has been helpful. “I feel like the process helps me to focus on areas in need of improvement” (Principal 35). Another participant explained this focus and how the system keeps the focus on the right things, “It will keep me focused on what I need to be doing as a principal. It keeps the important aspect of the job in front of my mind rather than getting so focused on the day to day grind” (Principal 4). These participants appreciated the focus on their work provided by the system.

Focus also referred to specific areas of focus. This sub-category came from participants specifically naming areas they were focusing on because of the CMES. Many mentioned a re-focus on being an instructional leader. One participant stated “Given me clear direction to base my goals and work habits toward to focus on improving my role as the facilitator of learning in my building” (Principal 50). Some mentioned a new focus on the school community, building relationships with parents, et
cetera. An comment supporting this was, “Through my own self-assessment, reflection and goal setting, I have used the CMES to ensure that the work of myself and my team is tied to the student performance and teacher growth; along with working with our school community” (Principal 7) This demonstrated the participant re-focusing on the priorities set by those who created the rubric/ system.

**Improvement.** Improvement as a category relates the sample principals wanting to get better at being a principal. As was stated earlier, this theme often was coupled with either focus or reflection. Improvement was mentioned as a category by 30 of the participants.

Many of the participants found that the feedback according to the CMES allows them to focus on areas in which to improve. Principal 66 stated “The evaluation process has provided me opportunities to identify areas of my job that I need to improve upon and allows opportunity to discuss ways to improve with my evaluator.” One participant even referenced how deficiencies shown by low ratings in the rubric can be worked on, “By knowing which areas I score the lowest, I am able to focus more time and energy on those areas once I understand what they are asking for and mean” (Principal 65).

Participants also mentioned that the feedback on the system gave them opportunities to focus on improving their schools. Principal 82 stated “Develop my next steps in school improvement” as the reason behind the change in practice. According to this sample, the CMES facilitates the focus of principals on improvement.

Improvement also related to the feedback from CMES has impacted practice because principals believe it has helped them try to improve their practice. Principal 47 stated, “I have been able to prioritize my professional growth.” Principal 86 stated “It
has opened the door for more in-depth conversations guiding and supporting documents and measures allow an individual to determine their own professional growth.”

Participants who mentioned this focused on the intent by those who created the CMES to improve their professional practice.

Although many of the participants mentioned improvement and focus in their comments, improvement needed to be a separate theme for this question. These participants not only appreciated that the CMES impacted their practice by that which they needed to focus, but also required them to focus on improving. “I always welcome opportunities to improve my practice” (Principal 5) shows one participant’s willingness to improve.

**Evaluator dependent comments.** This category related to the impact evaluators have on the CMES evaluation process. Comments coded under this category occurred 21 times over the 86 answers. This category manifests itself as supervisors driving the process or not driving the process.

Evaluator dependent comments includes comments from supervisors that have guided the focus of the sample participants in order to help the principals’ practice. In some of the comments, participants actually wrote about how their evaluators provided things on which principals were supposed to focus. From the evaluators comes the focus of these participants’ work. One strong example of this is, “When the superintendent evaluates me based on what is seen, heard, dealt with on a weekly, monthly, basis, there are focus areas that she will tell me that I need to work on so the scores will go up” (Principal 27). Another participant provided an example of informative feedback, “Because I am in my first year in my current position, I have had very informative and
substantive meetings with my evaluator” (Principal 78). Participants spoke of their feedback being changed because of their evaluator.

In addition, the participants referenced evaluators whose comments did not impact principals’ practices. Some of the participants mentioned that they have improved on their own, without effective evaluator feedback. “The feedback from my evaluator has not changed my work however my use of the system has changed my work” (Principal 11). Others spoke of the feedback they have received and it having minimal impact. “I don’t get as much feedback as I would like, I provide myself with areas to grow” (Principal 16). These participants believed that the feedback they receive has not impacted their practice.

Reflection. The last category for those participants who believed that the CMES has positively changed their practice is reflection. This category can be defined as the principals believe their practice is impacted by their ability to reflect on what they are doing or wanting to accomplish. Sixteen of the 86 principals had comments that were organized under this category.

The self-assessment is the tool that is used for principals to reflect on where they believe they are doing well and areas of growth within their practice (CDE, 2016). Those participants who mentioned the self-assessment or self-reflection mentioned it as a tool to identify where to grow. One participant stated “After my self-assessment I have identified areas of growth for myself resulting in more explicit work in the different areas in the coming year” (Principal 11). The self-assessment is the original way for principals to reflect.
Those participants who mentioned reflection wrote of how the system allows for them to reflect on their jobs as principals. These participants either mentioned reflection specifically or used words like think about or more conscious of to describe how they reflect on their practice. One participant stated “It has changed my work because it has made me highly reflective on my practices…” (Principal 13). Another wrote about the whole system being reflective, “I like that it encourages me to think about and set professional goals each year” (Principal 26).

**Feedback from the Colorado Model Evaluation System has had no impact.**

There were 52 participants who stated the CMES did not have any impact on their practice. There were three categories that emerged from coding the comments of these 52 participants: evaluator dependent, cumbersome system, and I am who I am. Those three themes are in order from largest number to smallest number.

*Evaluator dependent.* The first category found from the answers in this question was that there were some evaluator dependent variables. This category can be defined as the evaluator being the primary person who gives evaluative feedback and that this feedback was not impacting practice according to CMES. Twenty-seven comments were coded and sorted into this category. This category suggests that participants’ experiences with the CMES evaluation can depend on who they have as an evaluator.

Participants observed that the feedback was good but not attached to CMES. A couple participants mentioned that they have always received good feedback from their evaluators, but that CMES has not changed this. Principal 31 provided a good explanation of this by saying: “Superintendent has always been good at providing feedback, don’t feel like the CMES is why.” Another participant further explains this
perspective, “The feedback has not changed as a result of the system, I would receive the same feedback no matter the system with my evaluator” (Principal 51). One of the principals wrote about the ability of a supervisor to give more timely feedback outside of the system.

   My feedback has been more practical based on my situation and my learning needs-it was not specific to the CMES. Conversations were based on the reality of my school, teachers and students not a list on the rubric. It was those conversations that helped me grow and learn (Principal 37).

According to these participants, impactful feedback has occurred no matter the system, or outside of the system.

   Participants also mentioned that the feedback was, at times, ineffective. The participants were concerned about either receiving no or bad feedback from their evaluator, or that their evaluator was not adequately able to connect the feedback to their performance. The points made by these participants had to do with evaluators inability to connect feedback to either improving performance or to the CMES system due to lack of skill or training. One strong statement to this was, “I have not received any feedback that has detailed what specifically I need in order to improve my practice. The form has not led to my supervisors being able to use it to help guide how they will provide professional development for me to improve” (Principal 6). Principals also talked about not receiving enough feedback. This was explained by Principal 36 “No feedback.”

   I am who I am. The second category that became apparent in the answers to this question was that the principals responded by saying that they were not going to change how they approach their work because of an evaluation system. This category was
mentioned 14 times over the 52 answers. Some principals felt that they were going to strive to get better no matter the system. Principal 9 wrote, “I strive to meet the standards of my profession, regardless of the instrument which is used.” Others in the sample named those parts of their job which give the feedback they use to improve. One participant stated, “The changes I have made in my work over the years is based on self-reflection; parent, teacher, student perceptions; and conversations with colleagues from across the state” (Principal 12). Another wrote, “I have always been goal oriented and a hard worker focused on running a good school. This system made evaluation less objective and more focused on the roles of the principal, but has not changed my dedication or drive to be the best that I can be” (Principal 42). The statements written that were put into this category were about the hard work principals do no matter how they are evaluated.

_Cumbersome._ The third category for the participant answers to the final research question “no impact” is that the system has made things more difficult for people who use the process. Of the 52 total comments given, 12 of them fit under this category. According to the researcher, there are two sub-categories that best support the main category: a hoop to jump through and cumbersome system.

Half of the participants whose answers fit this category cited that the CMES is a bureaucratic system. This made the feedback, and the system, another hoop to jump through. One participant stated “The current evaluation system is another required activity that we simply endure and try to “get through” as quickly as possible” (Principal 18). Another participant dismissed the CMES by typing, “I don’t evaluate my performance based on the CMES. I would if it was not such a bureaucratic document”
This group of participants felt like the system was another hoop to jump through.

The other half of the participants who gave answers that fit this category were frustrated by the system being so cumbersome. “[CMES] is too cumbersome and slow of a process to provide meaningful feedback, evaluation, or assistance of any kind in the ‘real world’ of day to day school management and leadership” (Principal 8). Others described added burden to their (or their evaluator’s) workload. One mentioned, “This evaluation tool has not changed how I do my job. It simply has added extra work to my workload” (Principal 32). The evaluation system was seen as inefficient and burdensome.

*Feedback from Colorado Model Evaluation System has had a negative impact.*

The amount of those principals who have been impacted negatively based on feedback in the CMES system number three of the total sample. One of the three responses was a principal who had a bad experience with a recent evaluator, and wrote about that in their qualitative answer. The second wrote about the teacher evaluation instrument. Principal 3 wrote, “I now have less time to be an instructional leader. Also, I am not sure the model has had the impact intended.” All three had different reasons for their negative impression of the CMES.

*Conclusion to research question number 3.* This research question had a mixed analysis between the quantitative question and the qualitative question which allowed for the researcher to sort the qualitative questions by the answers given to the quantitative question. The median and mode for the quantitative question were both “positive”. The researcher found that 60.8 percent of the participants had a positive or extremely positive
change to their practice based on the feedback given using the CMES. The following categories were determined from the participants’ who had a positive impression’s answers: focus, improvement, reflection. There also was a large percent of the participants who felt the CMES had not impacted their practice. From this sub-sample of participants, the following categories were determined from their comments: evaluator dependent, I am who I am, and cumbersome. There were only three participants who answered with a belief that the feedback from CMES had a negative or extremely negative impact on their practice. The answers to research question #3 centered around feedback having a positive impact or no impact, with people who answered with “positive impact” or “no impact” giving answers with very different themes.

**Conclusion to Chapter IV**

This chapter was broken into sections according to the order of the survey questionnaire that was given to the participants. The demographics of the section were detailed. Each research question had both quantitative and qualitative data to analyze and compare. According to the results of the survey questionnaire the participants had a positive or neutral impression of the subjects all three research questions. After analyzing all of the results from the qualitative questions, distinct categories became apparent. In Chapter V, the researcher will provide further analysis of the three research questions and a general analysis of the study as a whole.

---

50 Percent positive for all three research questions was Q1 – 50.3, Q2 – 56.3, Q3 – 60.7. Percent neutral for all three research questions was Q1 – 40.9, Q2 – 41.0, Q3 – 37.1.  
51 Qualitative themes for all three research questions can be found in Appendix I.
CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

After analyzing the data, the researcher revisited the stated purpose of the Colorado Model Evaluation System (CMES), as identified by the State Council for Educator Effectiveness (2011), which is, “to provide meaningful and credible feedback that improves performance” (p. 7). The focus of this examination was to measure principals’ perceptions of CMES and determine if this evaluative tool is achieving its goal: “Today’s educators need and deserve an evaluation process that is designed to drive substantive feedback and vigorous conversations about effective teaching and effective school leadership” (State Council for Educator Effectiveness, 2011, p. 32). In this chapter, the researcher will explain the potential implications of the findings of the participants answers to the following three research questions:

Q1 What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process?

Q2 What are principals’ perceptions of the CMES evaluation feedback?

Q3 According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

In Chapter V there are six sections. After the introduction, final thoughts on the low response rate are given by the researcher. Next, the findings, which were detailed in Chapter IV, are coupled with the research from other researchers to provide context for the conclusions. Then, a synopsis of the limitations of this research study is presented.
The second to last section includes the implications on further research. This will be followed by the conclusion to Chapter V.

Response Rate

As was highlighted in Chapter IV, the response rate was less than expected at 16.02 percent. The concerns around this response rate and the potential bias of those who took the survey will be explained in the limitations section. The response rate for this study fit into the range provided by Baruch and Holtom (2008) for research surveys given to individuals in 2005 and approximated the low end of the range provided by Sivo et al. (2006). The researcher believes the method (survey questionnaire) and effort put into getting survey responses was sufficient. 52 By sending the questionnaire out to 951 participants and with the varied demographics of participants who took the survey questionnaire, the researcher also tried to ensure maximum variation in the sample. With these reasons in mind, the researcher decided to move forward with analyzing the data of this mixed methods study.

Discussion of the Three Research Questions

The Likert responses to the first two research questions provided context for the general impression the participants had of CMES and the feedback they received. For the third research question, the information provided showed the general impression of the participants and allowed for the answers of the sample to be sorted according to whether they had a positive, neutral, or negative connection to the question. The qualitative

52 The researcher chose principals who did not respond to the original survey to do a follow up non-responder survey. I only had one person of the 15 respond to my original solicitation. This person then did not respond to my request for an interview or a shortened survey questionnaire.
questions that followed the quantitative questions provided themes of both the strengths and challenges of the evaluation system and feedback according to the participants.

For all three research questions, over fifty percent of the participants gave a positive answer to the questions with Likert Scale answers. Those who gave neutral answers to the research questions were between thirty-seven and forty-one percent. As a result, it appears participants either have had a positive or neutral perception of the feedback used in conjunction with the CMES.

**What is Going Well?**

After analyzing the data in Chapter IV there were two major categories that spoke to what was going well with the feedback that is a part of the CMES: focus and improvement of participants’ practice. With focus, the participants identified how different parts of the system provided direction for them, whether it was the rubric (self-assessment), the goal setting, the feedback from evaluators, or the system structure requiring participants to focus on their evaluation. Focus was derived from the principal standards and the rubric. Kluger and DeNisi (1996) identified focus (on standards or goals) as an important step in Feedback Intervention Theory (FIT) “Behavior is regulated by comparisons of feedback to goals or standards” (p. 296). CMES, based on the participants’ responses in this study, enables administrators to link their feedback to the stated goals or standards for the state.

Some participants also appreciated that the CMES is based on improvement of practice. This major category tied in the goal setting, the system structure, and that improvement was a part of the expectations of the participants. With the CMES system being an evaluation of principals’ practices, having feedback within the system that helps
provide focus and is dedicated to improvement of practice was evident in the overall themes of these sample participants. With improvement as a theme, the goal of the CMES “to provide honest and fair assessments about educator performance and meaningful opportunities to improve” seemed to be somewhat successful with the participants in this study (State Council for Educator Effectiveness, 2011, p. 7).

**What Needs to Improve?**

Two major categories came to light about how feedback as a part of the CMES can improve when analyzing all the answers. These two major categories were: evaluator alignment to the system and making the system less of a burden. Although there were many differentiated thoughts and suggestions that fit under each major category, improvements to evaluators interaction with the system and to the system as a whole were identified throughout this study.

Brockner and Higgins (2001) named feedback as one specific way people of authority in an organization can impact those who work for them. The category centered on the role of the evaluator demonstrated the impact evaluators had on the perception of the participants. There were comments from the participants in the sample about evaluators that were: favorable to the evaluator, favorable to the system, negative about both, and/or critiques about specific ways evaluators interact with the sample participants. In the research sample, some participants thought highly of their evaluator and not the system. Some participants had negative experiences with their evaluators, but still felt the system was effective. Some participants had negative experiences with both. Superville (2016) recommended a prioritization on the improvements of principals and suggested the evaluation process to align evaluators to the practices of principals.
Participants’ comments demonstrated a disconnection between evaluator, system, and those being evaluated. In this study, it was also found that those participants who claimed to be receiving feedback once a semester or once a year were less likely to give positive answers to the closed-ended research questions. This would support Reeves’ (2004) findings, “A better model would provide specific, accurate, and timely feedback. Rather than an event that occurs once a year, evaluation should consist of frequent feedback and provide multiple opportunities for continuous improvement” (p. 57).

Evaluator alignment to the tool will be critical to the success of the CMES system.

The second major category (of potential improvements) that was prevalent throughout was clarifying or streamlining the system as a whole, with the majority of these thoughts relating to the rubric and the elements and indicators within the rubric. Kluger and DeNisi (1996) argued that “Attention is normally directed to a moderate level of the hierarchy” (p. 259). Also, attention should not be directed “to the ultimate goals of the self or to the detailed components of an ongoing activity” (Kluger & DeNisi, 1996, p. 262). Too many (or redundant) indicators might have principals focused too much on the details. Unclear indicators or elements might cause principals to focus too much on themselves and not their performance. With almost forty percent of the participants in this research stating that the feedback received as a part of the CMES either not impacting or negatively impacting their practice, clarifying or streamlining the instrument or system, may improve the principals’ experience with the CMES.

Thirty-seven percent of participants in this study who answered the third research question chose that the CMES has not changed their work as a principal. When asked to clarify their position, the answers of the participants gave reasons behind not engaging with the system: the evaluator, the system, or belief that the participants will try to do their job well and do not depend on an evaluation system. Those whose comments fit under evaluator dependent demonstrated that their perception of the impact of feedback depended on the evaluator, either blaming the evaluator for the experience, praising the evaluator with negative comments about the system, or blaming both. Participants were also frustrated with the system being cumbersome because it was seen as a bureaucratic system that took up too much time and resources to complete. This would connect with researchers who have critiques of evaluation systems. “Problems are seen with evaluation systems that have poorly stated criteria and standards of performance that leave questions in principals' minds about expectations” (Ginsberg & Berry, 1990, p. 220). This category would also contradict what Murphy (2003) stated as the attempted shift by those who created the Interstate School Leaders Licensure Consortium principals’ standards from education being a part of a bureaucracy to a system that is built to support learning. For the participants who believed that feedback using CMES had no impact, it can be concluded that the perceived lack of impact was because the feedback given (or the system) was viewed as something that did not relate to their job duties, added to their already burdened job, or was being done to them because of bureaucratic requirements (not because it was to improve their performance).
Research Question Number 2 Approaching Statistically Significant.

One calculation of ANOVA came back as approaching statistically significant. The school size demographics was approaching significance at .06. When the researcher looked further into this data, the difference between the schools with 1-300 students was not significantly different than those with 301-500 students, but was significantly different than schools with 501+ students. For this sample, the participants might have a different experience with feedback as a part of the CMES according to the size of their schools.

Implications

The second assumption by the State Council for Educator Effectiveness (the Council) in its Report and Recommendations (2011) was “the implementation and assessment of the evaluation system must embody continuous improvement” (p. 6). As a part of this process, the State Council for Educator Effectiveness (2011) also recommended that the state of Colorado monitor and act on “emerging research and best practice findings with respect to educator evaluations” (p. 6). The hope is that this research helps to inform the state of Colorado as it performs its cycle of continuous improvement. The implications of this research are further explored in the following paragraphs.

With the majority of the participants in this research study having a positive impression of the feedback that participants were receiving as a part of the CMES and also believing that feedback because of this system was positively impacting their practice, the system as a whole seems to be working for the majority of participants. By way of reminder, the stated purpose is, “to provide meaningful and credible feedback that
improves performance” (State Council for Educator Effectiveness, 2011, p. 7). With improvement being a major category in this study and most participants having a positive impression, then the purpose seems to be coming to fruition for a majority of the participants.

In the qualitative parts of this study, participants seemed to really appreciate that the system provided them a way to focus on their practice to improve based on a common set of best practices. This aligns with the findings of Kluger and DeNisi (1996) related to effective feedback interventions. This also aligns with researchers of principal evaluations who stated that effective evaluations have clear goals and standards (Stronge, 2013; Ginsberg & Berry, 1990; Davis et al., 2011). The researcher also believes that the focus provided by the system and the fact that the participants appreciate the CMES’s emphasis on improvement helps evaluators provide feedback that can be impactful on principal’s improvement.

Continuous improvement also considers what can be improved, and the results of this research show some parts of the CMES system that could be improved, according to the perspectives of the participants. There were evaluator specific suggestions. There were system specific concerns. For the third research question, 39.2 percent of the participants who responded to this question stated that feedback used as a part of the CMES had either no impact or a negative impact on their work as a principal.

With evaluators being those who give the feedback based on the CMES, having discrepancies in the quality of feedback and potentially in the perceptions of those giving the feedback could be issues with effective implementation of the evaluation system. If the standards that the CMES rubric are an effective demonstration of best practices for
principals, effective feedback to help support principals’ alignment with these standards is imperative. The relationship between evaluator and principals to provide effective feedback also impacts principals’ perceptions (Davis et al., 2011). Future research and discourse along the lines of how evaluators are trained to help principals with their job, could give a strong foundation for how to effectively implement Council of Chief State School Officers’ (2015) suggested standards to help evaluators best support and develop building leaders. Hattie (2012), when referencing feedback given to students by teachers stated, “The aim is to provide feedback that is ‘just in time’, ‘just for me’, ‘just for where I am in my learning process’” (p. 122). It would seem reasonable that principals would receive the same level of feedback from their evaluators. Future research can investigate ensuring that principal supervisors/evaluators receive the necessary training to ensure effective feedback is given.

For the 2017-2018 school year, CDE made changes to the rubric that narrowed the number of standards, elements, and indicators for the teacher evaluation, and piloted the new teacher rubric with districts throughout the state (CDE, nda). It would help the principal evaluation if CDE was to “make a similar effort as was made on the teacher rubric to make it shorter, more practical to use” (Principal 6). Participants in this study mentioned the amount of time the evaluation took, how many redundant or extraneous parts of the rubric there is, and how cumbersome the system was all. If some effort could be made on the principal’s rubric, this would create a shorter and more efficient system for principals and their evaluators.

Almost forty percent of the participants stated that their practice was not impacted or negatively impacted by the feedback according to CMES. The themes from those who
answered the third research question were evaluator dependent, cumbersome, and I am who I am. The first two suggestions in this Implications section would hopefully help remedy the issues brought to light by those of the 52 participants who felt that the feedback given as a part of the CMES has no impact on their practice. Further research needs to be done on those principals who might feel “I am committed to my work regardless of the evaluation” (Principal 10). Research could investigate how evaluation systems might be created or perfected that might authentically fulfill the intended purpose of the CMES as proposed by the Council to “provide meaningful and credible feedback that improves performance” (State Council for Educator Effectiveness, 2011, p. 7).

**Limitations of the Study**

Creswell (2012) identified the limitations of a study as “potential weaknesses or problems with the study identified by the researcher” (p. 199). There were many potential weaknesses or problems with this study. These limitations came from using a survey questionnaire, not having access to all principals, the use of a computer survey distribution and collection system, the inconsistencies and unknown errors of the purchased email list, the use of assistant principals in the pilot, the use of non-probability sampling, and the bias of those who take surveys.

The first limitation came when the researcher decided to use a survey questionnaire. The participants’ voices and stories were limited to their answers on the survey. “It is difficult to come to a deeper understanding of processes and contextual differences through questionnaires, which are standardized by nature and by their nature limited in length and depth of responses” (Muijs, 2012, p.141). This especially came true with some of the participants giving answers to the qualitative questions that ranged from
one or two words to one or two sentences. There was only one answer that was longer than four sentences. This was taken in consideration when the researcher decided to try online surveys instead of interviews, and the researcher decided the attempt to get a lot of perspectives outweighed this potential limitation.

The second limitation might be the choice to use assistant principals being evaluated with CMES as the sample population for the pilot study. Muijs (2012) recommend that “you need to sample a small group of respondents from the population you are interested in” (p. 153). The final population that the sample of participants for this study was drawn from was all principals who were being evaluated using CMES.

Assistant principals, though evaluated using the exact same rubric and system, are not principals and therefore not a part of the target population.

The limited scope of this study, which focused on the perceptions of principals in Colorado who were evaluated using the CMES, was another limitation. Those who developed the CMES system allowed school districts to forgo the CMES to build their own system. The school districts who created their own evaluation system only had to use the CMES as a model (State Council for Educator Effectiveness, 2011). Many of those school districts that created their own evaluation system are some of the largest, based off student population, in the state (CDE, 2016). There were a number of larger public-school districts that required the researcher to go through an Institutional Review Board (IRB) process prior to allowing their principals to take the survey. This action eliminated another potential pool of principals from the study. This study did not have access to a large number of principals in the state of Colorado.
Another set of limitations came when the researcher decided to purchase a list of principals from a recommended website. This purchase limits the scope of the list to only those principals whose email addresses were collected by the website. After having to go through and delete many principals’ emails because of the reasons mentioned in the previous paragraph, there were a few email addresses of principals found who were known to not have those email addresses anymore.

A final set of limitations for this study is the potential for bias; especially for those who tend to respond to survey questionnaires. “People who volunteer to be a part of a survey are often untypical” (Muijs, 2012, p. 145). Muijs (2012) stated that people who volunteer tend to have strong opinions, tend to have a lot of time on their hands, and also tend to have strong views about the research being done. Based on this finding from Muijs (2012), those that completed the survey may have only been principals with strong positive or negative opinions of CMES. The findings in this study may not reflect the views of all Colorado principals.

For Likert scale responses, Darity (2008) highlighted multiple biases that might prevent people from self-disclosing, “Ability to self-disclose is affected by such factors as reading comprehension (of the scale’s statements, instructions, etc.), experience, and familiarity with the subject matter of the scale” (p. 447). Motivation or willingness to accurately self-disclose can also be an issue. “Respondents may be motivated by any number of reasons to understate or to overstate their attitudes” (Darity, 2008, p. 447). Optimizing the ability of a person to accurately self-disclose is very important to the success of the Likert responses. All these potential biases are possible limitations.
With these limitations in mind, and with the low response rate, the final limitation of this study is that it is not a representative sample of the population of principals being evaluated. Sufficient effort was made to gather survey questionnaires. Most demographics were comparable between the participants and the data provided by CDE. The researcher used non-probability sampling, purchased a list with potential flaws, and had a 16.02 percent return rate on the final data used for the study. This sample of participants is not representative of the population.

**Recommendations for Future Research**

Overall, many of the participants in this study felt the CMES system and the feedback received because of this system positively provided them a strong focus that is based on improvement. According to the participants in this study, the CMES system could improve through better evaluator alignment and the streamlining/clarifying of the system, rubric, elements, and indicators within the rubric. Further research needs to be done to see what effective principal and evaluator interactions look like in the feedback and evaluation process. This could be done specifically with the CMES to improve the feedback within the CMES.

Further research also needs to be done on what effective evaluation/feedback systems look like that are not perceived to be a burden because of the amount of work. How does evaluation balance thoroughly accounting for that which is a principal’s job without naming every action of the principal?

Although the largest number of principals in the sample population had a positive impression of feedback, there were many who responded with neutral answers. Within the qualitative questions requesting positive answers, there were also many negative
comments about the CMES system and feedback given as a part of this system. This also would be a recommendation for future research. If there are a significant number of participants who have neutral or negative feelings towards the CMES, or feel the system has no impact on their practice, how can evaluation systems be crafted to shift this population from a feeling of neutral or negative to a positive impression of evaluation?

Finally, further research needs to be done on what effective feedback looks like between principals’ evaluators and principals themselves within an effective and efficient evaluation system. Ultimately, how do people who create evaluation systems best follow through with: “The purpose of the system is to provide meaningful and credible feedback that improves performance”? (State Council for Educator Effectiveness, 2011, p. 7).

**Recommendations for the Colorado Department of Education**

A second set of recommendations were created by the researcher for the Colorado Department of Education after the data was analyzed for these 152 participants. The recommendations fit into two major categories. These recommendations are suggestions for the system and for principal evaluators.

**System Based Recommendations**

Davis et al. (2011) stated, “Principal evaluation systems appear to be most effective when they are based upon clear standards and expectations of performance and aligned with the key goals and needs of principals, schools, and districts” (p. 35). First, some efforts need to be made to clarify and streamline the rubric. The pilot for the teacher rubric has been mentioned multiple times throughout this study. The researcher for this study is currently a part of this pilot, and the amount of time spent on the rubric is less and the feedback and coaching is no less impactful. Second, and this is already
happening in some capacity, continue to reach out to principals and their evaluators to find the most efficient ways to engage with the evaluation system. CDE has sent surveys that ask about principals’ and teachers’ experiences with the evaluation and feedback as a part of the evaluation. A next step in this process would be to have those who are being evaluated as a part of the CMES be asked in person (either by interview or through focus groups) to provide context to their perspectives. It would be advisable to create a system of continuous improvement similar to a professional learning committee in a school, where the practitioners are sharing perspectives and solutions. Third, if the frequency of feedback for the whole state of Colorado is similar to the participants in this study, then CDE needs to look into: How can the frequency of feedback increase without the burden of the system increasing? Potentially, with these suggestions, a counter narrative can be created to, “The process is time consuming at times and at other times feels like a compliance box to check off. The system should feel natural and be more responsive to the individual, allowing more flexibility” (Principal 48).

**Evaluator Based Recommendations**

The interaction between principals and their evaluators is extremely important. Davis et al. (2011) stated:

> The quality of the conduct of principal evaluation may be more important than its content. Strong, trusting, and collaborative relationships between principals and their district office evaluators is especially critical to the success of the evaluation process. (p. 35)

The following recommendations would be in support of improving the conduct of the principal evaluation to ensure its success. To improve the conduct of the evaluation
process, three steps of training principals’ evaluators are suggested. First step, would be
a development of understanding of how the evaluation process exists when functioning
properly. This type of training or dialogue would just be short training on the tool: how
the systems is supposed to work, how the rubric is supposed to work, how the feedback
within the system is supposed to work. Second step in this process would be the
development of the baseline understanding of all of the elements and indicators within
the rubric. Because of the number of current indicators, the training around this piece
could entail either the least understood indicators and elements, or a set of elements that
are the high impact elements for the rest of the rubric. Finally, a concerted effort should
be made on engaging those who are principal evaluators in training to more effectively
deliver feedback to principals that impacts principal practice. As an evaluator of
teachers, the researcher for this study aligns his methods of coaching around the model
provided by Cognitive Coaching. Regardless of the model, principal evaluators need to
be able to incorporate effective methods of coaching principals to fulfill the purpose of
the third assumption of the State Council for Educator Effectiveness (2011) “The purpose
of the system is to provide meaningful and credible feedback that improves performance”
(p.7). CDE needs to help facilitate the system’s purpose.

**Conclusion**

The focus throughout this study was whether the purpose of the Colorado Model
Evaluation System for principals was being fulfilled according to the perceptions of those
who are being evaluated. Evaluative feedback to improve practice is a noble goal set by
the State Council for Educator Effectiveness (2011). This goal, when looking at the
positive responses for the closed questions and the positive qualitative categories of focus
and improvement, was relatively successful, according to the majority of the participants in this study.

The participants of this study also identified some significant points that need further research to make appropriate adjustments to the CMES. For the system to improve, better connections need to be made between evaluators, effective feedback for principals, and the CMES system. Finally, the participants felt the system as a whole, as well as some parts, needs to be made more efficient in order to become more effective.

The biggest question has to do with the significant percent of participants who answered that the CMES has had no impact or negative impact on their practice. Kluger and DeNisi (1996), in their final assumption about feedback interventions, stated, “FIs change the locus of attention and therefore affect behavior” (p. 259). Since the behaviors of principals are supposed to align with the standards, or the elements and indicators identified in the CMES, then the system should work to positively impact all principal’s behaviors.

A system of evaluation that continues to focus on its strengths and uses identified gaps in perception to help inform future work on the system will help this system. This researcher would recommend continued emphasis on improvement of principal practice and providing a focus on what to improve. Refining the instrument and better aligning evaluator practice may improve the participants’ experiences with the CMES system so that it is a positive influence on principals’ practice, especially for those principals who have not been impacted by the Colorado Model Evaluation System.
REFERENCES


Retrieved from

http://www.sos.state.co.us/CCR/DisplayRule.do?action=ruleinfo&ruleId=3101&deptID=4&agencyID=109&deptName=300&agencyName=301%20Colorado%20State%20Board%20of%20Education&seriesNum=1%20CCR%20301-87


http://www.ccsso.org/Documents/Analysis%20of%20Leadership%20Standards-
Final-070913-RGB.pdf

American Institutes for Research. Retrieved from
http://files.eric.ed.gov/fulltext/ED530748.pdf

https://doi.org/10.1016/j.econedurev.2011.09.001

Retrieved from
http://www.cde.state.co.us/educatoreffectiveness/statemodelevaluationsystem-
pilot

http://www.cde.state.co.us/educatoreffectiveness/2011-13principalsystempilot

Colorado Department of Education (1999-2013). *School View Data Center.* Retrieved from https://edx.cde.state.co.us/SchoolView/DataCenter/reports.jspx?_adf.ctrl-
state=pac20phbp_4&_afrWindowMode=0&_afrLoop=7520579653801277&_adf.
ctrl-state=6nomwp1xp_4
http://www.cde.state.co.us/educatoreffectiveness/rubric-for-colorado-principals

Colorado Department of Education (2016a). *State Model Evaluation for Principals/Assistant Principals.* Retrieved from
http://www.cde.state.co.us/educatoreffectiveness/smes-principal

Colorado Department of Education (2016b). *Summary of districts/BOCES using the State Model Educator Evaluation System, hybrid systems, or locally-developed systems.* Retrieved from
http://www.cde.state.co.us/educatoreffectiveness/assurancesreport_2016

Colorado Department of Education (2016c). *Pupil membership by district setting.* Retrieved from
https://edx.cde.state.co.us/SchoolView/DataCenter/reports.jspx?_adf_ctrl-state=pac20phbp_4&_afrWindowMode=0&_afrLoop=7520579653801277&_adf.ctrl-state=6nomwp1xp_4

http://www.cde.state.co.us/educatoreffectiveness/usersguide


https://advance.lexis.com/container?config=0345494EJAA5ZjE0MDIyYy1kNzZkLTRkNzktYTkxMS04YMJhNjBiNWUwYzYKAFBvZENhdGFsb2e4CaPI4cak6laXLCWylBO9&crid=36fda0c-0d82-43f4-995cf81643b47b39


ISBN: 9781416603139


APPENDIX A

FRAMEWORK FOR SYSTEM TO EVALUATE PRINCIPALS
(State Council for Educator Effectiveness, 2011, p. 95)
APPENDIX B

STATE SCORING FRAMEWORK MATRIX
FOR PRINCIPALS
**State Scoring Framework Matrix for Principals**

This matrix is only a SAMPLE to illustrate for the reader what form the final scoring matrix will take. The final matrix may have a different number of rows and/or columns; and actual placement of the ineffective, partially effective, effective and highly effective boxes will be recommended by the Council and CDE after analysis of the pilot data.

<table>
<thead>
<tr>
<th>Quality Standards Score</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partially Effective</td>
<td>Effective</td>
<td>Highly Effective</td>
<td>Highly Effective</td>
<td>Partially Effective</td>
<td>Ineffective</td>
</tr>
<tr>
<td>Ineffective</td>
<td>Effective</td>
<td>Effective</td>
<td>Effective</td>
<td>Partially Effective</td>
<td>Ineffective</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

(Student Growth Score)

(State Council for Educator Effectiveness, 2011, p. 124)
APPENDIX C

TABLE C.1: COLORADO DEPARTMENT OF EDUCATION’S SURVEY RESULTS OF PILOT PRINCIPALS’ PERCEPTIONS
### Table C.1

**Percent of Positive Responses Given by Principals in Baseline and Feedback Surveys**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Is intended to guide professional growth.</td>
<td>32</td>
<td>87</td>
<td>93</td>
<td>+6</td>
</tr>
<tr>
<td>Serves as a basis for improving teaching and learning.</td>
<td>13</td>
<td>60</td>
<td>70</td>
<td>+10</td>
</tr>
<tr>
<td>Is based on current scientifically sound research.</td>
<td>18</td>
<td>59</td>
<td>83</td>
<td>+24</td>
</tr>
<tr>
<td>Provides actionable feedback to the person being evaluated.</td>
<td>26</td>
<td>66</td>
<td>75</td>
<td>+9</td>
</tr>
<tr>
<td>Sets high standards for the person being evaluated.</td>
<td>43</td>
<td>81</td>
<td>88</td>
<td>+7</td>
</tr>
<tr>
<td>Supports the improvement of the school’s instructional program.</td>
<td>27</td>
<td>65</td>
<td>72</td>
<td>+7</td>
</tr>
<tr>
<td>Documents changes in professional practice over time.</td>
<td>20</td>
<td>57</td>
<td>75</td>
<td>+18</td>
</tr>
<tr>
<td>Identifies areas of strength.</td>
<td>48</td>
<td>81</td>
<td>92</td>
<td>+9</td>
</tr>
<tr>
<td>Identifies areas that need improvement.</td>
<td>48</td>
<td>78</td>
<td>89</td>
<td>+11</td>
</tr>
<tr>
<td>Provides an accurate assessment of my performance.</td>
<td>24</td>
<td>50</td>
<td>58</td>
<td>+26</td>
</tr>
<tr>
<td>Results in improved student growth.</td>
<td>19</td>
<td>37</td>
<td>44</td>
<td>+18</td>
</tr>
<tr>
<td>Identifies areas that need improvement.</td>
<td>48</td>
<td>78</td>
<td>89</td>
<td>+11</td>
</tr>
<tr>
<td>Provides an accurate assessment of my performance.</td>
<td>24</td>
<td>50</td>
<td>58</td>
<td>+26</td>
</tr>
<tr>
<td>Results in improved student growth.</td>
<td>19</td>
<td>37</td>
<td>44</td>
<td>+18</td>
</tr>
</tbody>
</table>

APPENDIX D

FINAL SURVEY FOR PRINCIPALS
Final Survey for Principals

Q1 This is a survey principals who are being evaluated under the Colorado State Model Evaluation System.

If your district created its own system, please do not take the survey. If you do not know whether or not your district is using CMES, click on the following link and scroll to pages 4-12. (If your district shows State/Green in the principal column, your district uses CMES. If it shows LOCAL or N/E then it DOES NOT)

Does my district use CMES?

You will be asked to answer demographics questions before answering the survey questions.

If you wish to review the purpose and components of the Colorado Model Evaluation System, please copy the following link into another window. (You could start on page 11)

http://www.cde.state.co.us/educatoreffectiveness/usersguide

If you wish to review the rubric used for the Colorado Model Evaluation System, please copy and paste the following URL to a separate window.

http://www.cde.state.co.us/educatoreffectiveness/rubric-for-colorado-principals

Please read through the consent form for human participation prior to taking the survey. You can find it at this link.

Consent for human participation

When piloted, the most frequent amount of time it took people to take the survey was 10 minutes, with the longest amount of time being 18 minutes.

If you have any follow up questions or concerns, please contact me.
(c) 970.999.2115
(email) ridd5584@bears.unco.edu

If you agree to participate in the following survey, please click the appropriate answer below. If you choose to not to participate in the survey, please choose the appropriate option below.

- I agree to answer the questions on the following survey and allow the researcher the opportunity to use my answers for research purposes.

- I DO NOT agree to participate in the survey, and choose to exit the survey as opposed to allow the researcher to use my answers for research purposes.
Q2 Your Gender

- Male
- Female
- I do not identify with a particular gender.

Q3 Years as a principal: Please type a number (1,2,3, etc.)

________________________________________________________________

Q4 Years in current position (1,2,3, etc.)

________________________________________________________________

Q5 Evaluator Type (Person who is assigned as your primary evaluator for the Colorado Model Evaluation System)

- Superintendent
- Assistant Superintendent
- Director of Secondary or Elementary Education
- Other ________________________________________________

Q6 Gender of Your Evaluator

- Male
- Female
- This person does not identify with a particular gender.
Q7 Type of District

- Urban
- Rural
- Suburban

Q8 Location of District

- Northern Mountains
- Central Mountains
- Southern Mountains
- Front Range
- Western Slope
- Eastern Plains
- Other ____________________________________________
Q9 Free and Reduced Lunch Rate of School

- 0-9%
- 10-19%
- 20-29%
- 30-39%
- 40-49%
- 50-59%
- 60-69%
- 70-79%
- 80-89%
- 90-100%

Q10 Size of School (2016-2017 School Year)

- 1-100 Students
- 101-300 Students
- 301-500 Students
- 501-700 Students
- 701-900 Students
- 901-1100 Students
- 1101-1300 Students
- 1301-1500 Students
- 1501-1700 Students
- 1701-1900 Students
Q11 Type of School

- Elementary (K-5)
- Middle (5-8, 6-8)
- Junior High
- High School
- K-8
- K-12
- Other ________________________________________________
Q12 What was your Final Effectiveness Rating for the 2016-2017 School Year?

- Highly Effective
- Effective
- Partially Effective
- Ineffective
- I was evaluated, but not using the CMES last year
- I was not evaluated as a principal last year.

Q13 How often do you receive evaluative feedback for the Colorado Model Evaluation System?

- Once a day
- Once a week
- 2-3 times a month
- Once a Month
- Once every 2 months
- Once a Semester
- Once a year
- I have never received evaluative feedback in my current position.
Q14 What is your overall perception of the Colorado Model Evaluation System for Principals?

- [ ] Extremely Positive
- [ ] Positive
- [ ] Neither Positive Nor Negative
- [ ] Negative
- [ ] Extremely Negative

Q15 What do you like about CMES?

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

Q16 How do you think the CMES can improve?

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

Q17 What aspects of the CMES has been most helpful to you?

________________________________________________________________
________________________________________________________________
________________________________________________________________
Q18 What aspects of CMES have been least helpful?

________________________________________________________________

________________________________________________________________

________________________________________________________________

________________________________________________________________

________________________________________________________________

________________________________________________________________
Q19 What has been your perception of the feedback you have received as a part of the Colorado Model Evaluation System?

- [ ] Extremely Positive
- [ ] Positive
- [ ] Neither Positive nor Negative
- [ ] Negative
- [ ] Extremely Negative

Q20 How has the feedback you have received as a result of the CMES been helpful?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Q21 How could the feedback as a result of the CMES be improved?

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Q28 How has the feedback you have received using the Colorado Model Evaluation System changed your work as a principal?

- [ ] Greatly Changed (Positive)
- [ ] Changed (Positive)
- [ ] Has not changed my work.
- [ ] Changed (Negative)
- [ ] Greatly Changed (Negative)

Q22 Please thoroughly explain your answer to the previous question.

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________

Q23 Is there anything that you would like to add about the Colorado Model Evaluation System, that you were not able to express in the previous questions?

________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
________________________________________________________________
Q27 Thank you for your time and effort.

Click the button below and then exit the survey.

End of Block: Default Question Block
APPENDIX E

CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO
CONSENT FORM FOR HUMAN PARTICIPANTS IN RESEARCH
UNIVERSITY OF NORTHERN COLORADO

Project Title: According to Principals' Perceptions, how is the feedback from the Colorado Model Evaluation System evaluation process altering principals' practice?
Researcher: Tim Ridder M.A., Educational Leadership and Policy Studies
Phone Number: 970.999.2115 e-mail: ridd5584@bears.unco.edu
Research Advisor: Spencer C. Weiler Ph. D., Educational Leadership and Policy Studies
Phone Number: 970.351.1016 e-mail: spencer.weiler@unco.edu

As many districts implement the new Colorado State Model Evaluation System (CMES), many principals are being exposed to a new system of evaluation. These districts are supposed to “collect evidence of principal performance with enough frequency to ensure that principals are provided with ongoing feedback and the opportunity to improve performance” (State Council for Educator Effectiveness, 2011, p. 109). Feedback was so important to those who created the CMES that they stated: "The purpose of the system is to provide meaningful and credible feedback that improves performance" (State Council for Educator Effectiveness, 2011). Feedback being the purpose of the system, and the perception of the principals of this system are critical elements to the success of the system itself.

The survey will take you anywhere from 10-15 minutes to complete.

You will not be asked to submit any specific personal contact information on the survey. All of the initial questions on the survey will ask demographic information about you and your school. The collected information will be general enough so that you will not be able to be specifically identified as having answered the questions. The demographic information will be specific enough to create categories of information to help to see if trends occur. There is very little chance that your email-address will be connected to the demographics information that you provide. Your email will be attached to the survey only while the survey results are being tallied. These email addresses will be deleted prior to any analysis. If any connection is made, it will only be known by the researcher and will not be shared with anyone else. If you volunteer any information that specifically identifies anyone or any school in the open-ended section of the survey, the names will be altered to keep the answers anonymous.
After completing the demographic information, you will be asked to answer the questions on a 11-question survey. The survey will be sent to you through Qualtrics\textsuperscript{53}. Three of the questions are questions that will require Likert responses. The other eight questions will be open ended questions. All eleven questions relate to the three main research questions:

- What are principals' perceptions of the CMES evaluation process?
- What are principals' perceptions of the CMES evaluation feedback?
- According to principals' perceptions, how is the feedback from the CMES evaluation process altering principals' practice?

The risks to you are minimal. Qualtrics provides a survey format that can ensure the strictest security in not tracking those who have answered their survey instruments, and ensuring the security of information while storing the results of surveys.\textsuperscript{54} If there is a breach of security, Qualtrics has a policy and procedure that must be followed to solve the issue. You will be notified if security has been breached. Please notify the researcher by phone or email if you believe your information has been breached.

The discomfort to you is minimal. This survey should take you a minimal amount of time. If you are uncomfortable with being critical of the evaluation system in any way, please be assured that any answers you give will not be attached to your email address. The goal is to keep this survey anonymous.

The benefit to you by participating in this survey is that you are participating in new research regarding the Colorado Principal Model Evaluation System. Feedback is an important component to the improvement of principals in their craft, and this survey allows you the opportunity to comment on or rate the effectiveness of the feedback you are receiving.

Participation is voluntary. You may decide not to participate in this study and if you begin participation you may still decide to stop and withdraw at any time. Your decision will be respected and will not result in loss of benefits to which you are otherwise entitled. Having read the above and having had an opportunity to ask any questions please complete the survey if you would like to participate in this research. By completing the survey, you will give me permission for your participation. You may keep this form for future reference. If you have any concerns about your selection or treatment as a research participant, please contact Sherry May, IRB Administrator, Office of Sponsored Programs, Kepner Hall, University of Northern Colorado Greeley, CO 80639; 970-351-1910.

\textsuperscript{53} Qualtrics is an online survey instrument used by the University of Northern Colorado.

\textsuperscript{54} Qualtric’s security policy has been cleared by an IRB panel at the University of Northern Colorado.
APPENDIX F

UNIVERSITY OF NORTHERN COLORADO INSTITUTIONAL REVIEW BOARD APPROVAL
DATE: September 15, 2017

TO: Timothy Ridder, M.A.

FROM: University of Northern Colorado (UNCO) IRB

PROJECT TITLE: [707389-1] According to principals' perceptions, how is the feedback from the CMES evaluation process altering principals' practice?

SUBMISSION TYPE: New Project

ACTION: APPROVAL/VERIFICATION OF EXEMPT STATUS

DECISION DATE: September 15, 2017

EXPIRATION DATE: September 15, 2021

Thank you for your submission of New Project materials for this project. The University of Northern Colorado (UNCO) IRB approves this project and verifies its status as EXEMPT according to federal IRB regulations.

Tim –

Thank you for a clear and thorough IRB application for meaningful research. Please add Dr. Weiler's name and contact information to the heading of the consent form before use in participant recruitment and data collection. Otherwise, your application materials and protocols are verified/approved exempt.

Best wishes with your study.

Sincerely,

Dr. Megan Stellino, UNC IRB Co-Chair

We will retain a copy of this correspondence within our records for a duration of 4 years. If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within University of Northern Colorado (UNCO) IRB's records.
APPENDIX G
INITIAL EMAIL SENT TO ALL PRINCIPALS
Hello,

My name is Tim Ridder. I am a doctoral student at the University of Northern Colorado. I am researching principals' perceptions of the Colorado Model Evaluation System for Principals and the feedback you have received as a part of this system. I seek to find the answers to these three questions:

- What are principals’ perceptions of the Colorado Model Evaluation System (CMES) evaluation process?
- What are principals’ perceptions of the CMES evaluation feedback?
- According to principals’ perceptions, how is the feedback from the CMES evaluation process altering principals’ practice?

I would greatly appreciate if you would complete my survey sent to you on September 23, 2017.

**Initial Details:**

- Only those being evaluated using the CMES for Principals should take this survey. (I apologize if you receive this email and/or survey and are not being evaluated under CMES)
- To verify, click here: http://www.cde.state.co.us/educatoreffectiveness/assurancesreport_2016
  - Go to page 4
  - Look for your district
  - Under Principal - if it says 'State' then you should take the survey when sent on September 23, 2017
- The survey will be sent through a web based format called Qualtrics and appear to come from survey.reply@unco.edu (Tim Ridder).
- It should take you between 10 and 20 minutes. (The majority of pilot survey participants took 10 minutes or under)

Thank you for any and all time you spend supporting my research!

Tim Ridder
Doctoral Student
University of Northern Colorado
APPENDIX H

TABLES H.1, H.2, & H.3 – CROSSTABS COMPARING FREQUENCY OF FEEDBACK WITH THE THREE RESEARCH QUESTIONS
Table H.1

*Frequency of feedback using the CMES and principal perception of CMES – Crosstab*

<table>
<thead>
<tr>
<th>Frequency of Feedback</th>
<th>Extremely Positive</th>
<th>Positive</th>
<th>Neither Positive nor Negative</th>
<th>Negative</th>
<th>Extremely Negative</th>
<th>Total Responses per Row</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>n=1</td>
</tr>
<tr>
<td>Once a week</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=2</td>
</tr>
<tr>
<td>2-3 times a month</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>n=6</td>
</tr>
<tr>
<td>Once a month</td>
<td>1</td>
<td>10</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>n=19</td>
</tr>
<tr>
<td>Once every two months</td>
<td>0</td>
<td>11</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>n=17</td>
</tr>
<tr>
<td>Once a semester</td>
<td>0</td>
<td>31</td>
<td>23</td>
<td>2</td>
<td>1</td>
<td>n=57</td>
</tr>
<tr>
<td>Once a year</td>
<td>0</td>
<td>15</td>
<td>21</td>
<td>7</td>
<td>1</td>
<td>n=44</td>
</tr>
<tr>
<td>Never</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>n=3</td>
</tr>
<tr>
<td>Totals</td>
<td>3</td>
<td>72</td>
<td>61</td>
<td>11</td>
<td>2</td>
<td>n=149</td>
</tr>
</tbody>
</table>
Table H.2

*Frequency of feedback using the CMES and principal perception of feedback using CMES – Crosstab*

<table>
<thead>
<tr>
<th>Frequency of Feedback</th>
<th>Extremely Positive</th>
<th>Positive</th>
<th>Neither Positive nor Negative</th>
<th>Negative</th>
<th>Extremely Negative</th>
<th>Total per Row</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=0</td>
</tr>
<tr>
<td>Once a week</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=2</td>
</tr>
<tr>
<td>2-3 times a month</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=6</td>
</tr>
<tr>
<td>Once a month</td>
<td>2</td>
<td>10</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>n=20</td>
</tr>
<tr>
<td>Once every two months</td>
<td>1</td>
<td>9</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>n=17</td>
</tr>
<tr>
<td>Once a semester</td>
<td>2</td>
<td>33</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>n=53</td>
</tr>
<tr>
<td>Once a year</td>
<td>2</td>
<td>14</td>
<td>25</td>
<td>1</td>
<td>1</td>
<td>n=43</td>
</tr>
<tr>
<td>Never</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>n=3</td>
</tr>
<tr>
<td>Totals</td>
<td>8</td>
<td>73</td>
<td>59</td>
<td>3</td>
<td>1</td>
<td>n=144</td>
</tr>
</tbody>
</table>
Table H.3

Frequency of feedback using the CMES and principal perception of how CMES has changed principal practice – Crosstab

<table>
<thead>
<tr>
<th>Feedback Frequency for CMES</th>
<th>Frequency Greatly Changed (Positive)</th>
<th>Frequency Changed (positive)</th>
<th>Frequency Has not Changed my Work</th>
<th>Frequency Changed (Negative)</th>
<th>Frequency Greatly Changed (Negative)</th>
<th>Total Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once a day</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>n=1</td>
</tr>
<tr>
<td>Once a week</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=2</td>
</tr>
<tr>
<td>2-3 times a month</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>n=6</td>
</tr>
<tr>
<td>Once a month</td>
<td>1</td>
<td>11</td>
<td>6</td>
<td>4</td>
<td>0</td>
<td>n=19</td>
</tr>
<tr>
<td>Once every two months</td>
<td>0</td>
<td>10</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>n=16</td>
</tr>
<tr>
<td>Once a semester</td>
<td>1</td>
<td>34</td>
<td>15</td>
<td>0</td>
<td>2</td>
<td>n=52</td>
</tr>
<tr>
<td>Once a year</td>
<td>0</td>
<td>19</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>n=41</td>
</tr>
<tr>
<td>Never</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>n=3</td>
</tr>
<tr>
<td>Totals</td>
<td>4</td>
<td>81</td>
<td>52</td>
<td>1</td>
<td>2</td>
<td>n=140</td>
</tr>
</tbody>
</table>
APPENDIX I

TABLE I.1: QUALITATIVE CATEGORIES FOR ALL THREE RESEARCH QUESTIONS
### Table I.1

**Categories for all three research questions**

**Q#1 - What are the principals’ perceptions of the CMES evaluation process?**

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of comments that fit under this category</th>
<th>What can improve with CMES? (least helpful)</th>
<th>Number of comments that fit under this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expectations</td>
<td>75</td>
<td>Clarify/Streamline</td>
<td>111</td>
</tr>
<tr>
<td>Growth</td>
<td>56</td>
<td>System</td>
<td>41</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>28</td>
<td>My Job</td>
<td>16</td>
</tr>
<tr>
<td>Conversation</td>
<td>26</td>
<td>District Use of Tool</td>
<td>16</td>
</tr>
<tr>
<td>Negative Comments</td>
<td>13</td>
<td>Measures of Student Learning</td>
<td>14</td>
</tr>
<tr>
<td>Negative Comments</td>
<td>13</td>
<td>Disappear</td>
<td>4</td>
</tr>
</tbody>
</table>

**Q#2 - What are the principals’ perceptions of the CMES evaluative feedback?**

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of comments that fit under this category</th>
<th>How could feedback using CMES be improved?</th>
<th>Number of comments that fit under this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement</td>
<td>50</td>
<td>Specificity of feedback</td>
<td>27</td>
</tr>
<tr>
<td>District/Evaluator Specific</td>
<td>12</td>
<td>Frequency and timeliness of feedback</td>
<td>21</td>
</tr>
<tr>
<td>Positive Non-productive, not helpful, no feedback</td>
<td>10</td>
<td>Instrument and system based suggestions</td>
<td>17</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>55</td>
<td><strong>Total</strong></td>
<td>55</td>
</tr>
</tbody>
</table>

---

55 For the first research question, the themes were coded by the number of times the categories were mentioned. For the second and third research questions, the categories were coded by number of principals who mentioned the categories.
Q#3 - According to principals' perceptions, how is the feedback from the CMES process altering principal's practice?

<table>
<thead>
<tr>
<th>Feedback using CMES has positive impact</th>
<th>Number of comments that fit under this category</th>
<th>Feedback using CMES has No Impact</th>
<th>Number of comments that fit under this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus</td>
<td>47</td>
<td>Evaluator</td>
<td>Dependent</td>
</tr>
<tr>
<td>Improvement</td>
<td>30</td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>Reflection</td>
<td>16</td>
<td>I am Who I am</td>
<td>14</td>
</tr>
<tr>
<td>Evaluator Dependent</td>
<td>21</td>
<td>Cumbersome</td>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feedback using CMES has a negative impact</th>
<th>Number of comments that fit under this category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad Evaluator</td>
<td>1</td>
</tr>
<tr>
<td>Teacher Eval.</td>
<td>1</td>
</tr>
<tr>
<td>Answer</td>
<td></td>
</tr>
<tr>
<td>Less Time Consuming</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>