Mental Models Held by School Principals 
Regarding a New System of Evaluating Teachers

A dissertation submitted in partial fulfillment of the requirements for the 
degree of Doctor of Philosophy in Educational Leadership

by

Kirsten E. Gleissner

Dr. Bill Thornton/Dissertation Advisor

August, 2017
THE GRADUATE SCHOOL

We recommend that the dissertation prepared under our supervision by

KIRSTEN E. GLEISSNER

Entitled

Mental Models Held by School Principals Regarding a New System of Evaluating Teachers

be accepted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Dr. Bill Thornton, Advisor

Dr. Janet Ussinger, Committee Member

Dr. Thomas Harrison, Committee Member

Dr. Steven Pradere, Committee Member

Dr. Alex Kumjian, Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

August, 2017
Abstract

In recent years, interest has grown among legislators and educators regarding the use of teacher evaluation as a leverage point for increasing the effectiveness of instruction in schools. Though there is an existing body of knowledge around the attitudes and beliefs held by teachers about their own evaluation experiences, very little information exists about teacher evaluation from the perspective of the site administrator. The purpose of this qualitative study was to examine the mental models (beliefs, attitudes, assumptions, and knowledge) held by school principals regarding a new teacher evaluation system. The questions guiding the study were 1) What mental models (beliefs, assumptions, attitudes, and knowledge) do principals hold about the role of teacher evaluation as a vehicle for change in schools; and, 2) Do different school contexts (elementary, middle, and high school) influence the mental models that site administrators hold relative to teacher evaluation?

Sixteen principals were interviewed during the fall of 2016 regarding their perceptions of a new teacher evaluation system. The interviews were examined using a constructivist grounded theory methodology. Results of the analysis revealed three major themes with subthemes: 1) Tension: The Importance of Doing Teacher Evaluation and Feeling Overloaded and Overwhelmed, 2) Desperately Trying to Figure It Out: It’s a Catch-22 System and Getting a Handle on Time Management, and 3) Glimpses of Resolution. Findings in this study indicated that principals’ mental models play a significant role in the change process as related to school initiatives.
Acknowledgements

The completion of any major study depends on the quality and humanity of the dissertation committee and in this respect I was extremely lucky. I would like to thank the members of my committee for their time and support of my efforts: Dr. Thomas Harrison for his time and his trust in my work; Dr. Steven Pradere for juggling both district commitments and time for my committee; Dr. Alex Kumjian for his long-time friendship and interest in this endeavor; and Dr. Janet Usinger for her insightful, personalized guidance of my process and findings. To my advisor, Dr. Bill Thornton (Dr. T!): I am forever indebted to you for your patience, kindness, expectations, and expertise. You are a gem and, though it may sound cliché, you will know it is true when I say I could not have done this without your guidance and support!

I also wish to thank the principals who were willing to take time out of their extremely busy days to share their experiences, thoughts, and knowledge with me for this project. You have my greatest admiration and respect!

To my friends and family, words cannot express my gratitude for your caring and support through this process. To my amazing children, Kastania and Kristian: Thank you for your faith in me, your help, and your understanding of my distraction! To my marvelous sister, Vesta, thank you for listening, listening, and listening! To my dear friends, thank you for still being my friends! And to my fabulous and patient husband, Per Rasmussen: Tak fordi du altid har været der til at fange mig, knuse mig, og sparke mig – jeg elsker dig så højt og er her hvor jeg er udelukkende på grund af din støtte. I am sorry it took so long and I am looking forward to “life after doc” with you!
# Table of Contents

Abstract ................................................................................................................................................. i

Acknowledgements .......................................................................................................................... ii

List of Tables ........................................................................................................................................ vi

List of Figures ........................................................................................................................................ vii

Chapter One: Background, Statement of the Problem, and Research Questions.... 1
   Background ......................................................................................................................................... 1
   Statement of the Problem .................................................................................................................. 2
   Purpose of the Study .......................................................................................................................... 4
   Research Questions ........................................................................................................................... 4
   Significance of the Study ..................................................................................................................... 5
   Assumptions ........................................................................................................................................ 6
   Limitations ......................................................................................................................................... 6
   Ethical Considerations ....................................................................................................................... 7
   Definition of Terms ............................................................................................................................. 7
   Summary ............................................................................................................................................. 10

Chapter Two: Review of Relevant Literature ....................................................................................... 11
   Holonomy ........................................................................................................................................... 11
   Organizational Learning .................................................................................................................... 13
      Organizational Learning: A Process ............................................................................................... 14
      Single- and Double-Loop Learning ................................................................................................ 18
      Deuterolearning .............................................................................................................................. 22
   The Learning Organization: Structure ............................................................................................... 23
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Mastery</td>
<td>25</td>
</tr>
<tr>
<td>Mental Models</td>
<td>27</td>
</tr>
<tr>
<td>Building Shared Vision</td>
<td>28</td>
</tr>
<tr>
<td>Team Learning</td>
<td>29</td>
</tr>
<tr>
<td>Systems Thinking</td>
<td>30</td>
</tr>
<tr>
<td>Seven Organizational Learning Disabilities</td>
<td>33</td>
</tr>
<tr>
<td>Understanding Mental Models</td>
<td>35</td>
</tr>
<tr>
<td>Teacher Evaluation as an Educational Reform Initiative</td>
<td>41</td>
</tr>
<tr>
<td>Teacher Evaluation in Nevada</td>
<td>45</td>
</tr>
<tr>
<td>Chapter Three: Methodology</td>
<td>52</td>
</tr>
<tr>
<td>Research Design</td>
<td>52</td>
</tr>
<tr>
<td>Setting and Context</td>
<td>54</td>
</tr>
<tr>
<td>Schools</td>
<td>55</td>
</tr>
<tr>
<td>Participants</td>
<td>56</td>
</tr>
<tr>
<td>Data Source and Collection Procedures</td>
<td>58</td>
</tr>
<tr>
<td>Interview Process</td>
<td>60</td>
</tr>
<tr>
<td>Data Analysis Process</td>
<td>61</td>
</tr>
<tr>
<td>Role of the Researcher</td>
<td>63</td>
</tr>
<tr>
<td>Chapter Four: Results</td>
<td>65</td>
</tr>
<tr>
<td>Tension</td>
<td>65</td>
</tr>
<tr>
<td>Importance of Doing Teacher Evaluation</td>
<td>65</td>
</tr>
<tr>
<td>Feeling Overloaded and Overwhelmed</td>
<td>69</td>
</tr>
<tr>
<td>Desperately Trying to Figure It Out</td>
<td>78</td>
</tr>
</tbody>
</table>
List of Tables

1. Designation of Title I Schools................................................................. 56
List of Figures

1. Development of a Theory of Action ................................................................. 16
2. Single-Loop Learning ...................................................................................... 19
3. Single- and Double-Loop Learning ................................................................. 21
4. Senge’s Five Disciplines ............................................................................... 25
5. Nevada Educator Performance Framework Continuous Learning Cycle for Teachers and Site Administrators ......................................................... 49
6. Nevada Educator Performance Framework Teacher Standards ................. 50
7. Visual Representation of a Grounded Theory Model ...................................... 54
8. Benefits of the intensive interview approach for both the interviewer and the interviewee ............................................................................ 59
9. Continuum of Major Themes Identified in the Principal Interview Data Regarding Teacher Evaluation ................................................................. 94
10. Relationship between Major Findings .......................................................... 104
Chapter One: Background, Statement of the Problem, and Research Questions

As educators, legislators, and community members attempt to find ways to improve the education system, many different reform and change initiatives have been vetted to improve learning for students (Teachers and Leaders Council, 2013). Most recently, teacher evaluation has been identified as a leverage point to increasing the effectiveness of instruction (Daly & Kim, 2010; Leithwood, Louis, Anderson, & Wahlstrom, 2004; Mathers, Oliva, & Laine, 2008). According to the White Paper Record of Systems Development and Archival Record of the Deliberations, “educator evaluation serves as the foundation to increasing educator effectiveness and retention, and to equitable distribution of effective teachers and administrators” (Teachers and Leaders Council, 2013, p. 4).

Background

The State of Nevada began the formal process of developing a statewide educator evaluation system in 2011 when legislators passed Assembly Bill 222. With this new bill, the Teachers and Leaders Council (TLC) was established and given the task of researching and developing a statewide system of performance for teachers and school building administrators. The system was built on four levels of performance: highly effective, effective, minimally effective, or ineffective. Student achievement data was included in the rating system. The Nevada State Board of Education approved the system, based on the recommendations of the TLC, and the Nevada Educator Performance Framework, or the NEPF, was adopted (Nevada Department of Education, 2015). During the 2013 legislative session, Senate Bill 407 was passed; this bill further refined the NEPF to include, among other changes, a validation study of the system and guidelines...
for evaluation during the transition period to the new statewide system (Nevada Department of Education, 2015). The 78th Nevada Legislature of 2015 confirmed the use of the NEPF as the state educator evaluation system for teachers and administrators. With the passage of Assembly Bill 447, all educators in Nevada were evaluated using the same evaluation system for the 2015-2016 school year.

**Statement of the Problem**

In the past decade, the list of responsibilities of the school principal has become increasingly longer as accountability has become significantly more stringent (Balyer, 2014; Ediger, 2009; Ediger, 2014; Rayfield & Diamantes, 2004). Along with increasing managerial duties has been the added requirement of being an instructional leader. Not only is there more to do, but the role of school leaders has become intensified. According to Gronn (2003), “Intensification means that work becomes harder and harder to perform” (p. 65). Intensification of the work of school leaders in terms of teacher evaluation has been documented through federal grants and programs as they require teacher and administrator evaluations to have more influence in the approval of grants as well as the former No Child Left Behind flexibility waivers (United States Department of Education, 2015a).

Prior to 2015, district leaders in Nevada were able to choose their district’s evaluation system as well as the process of rotation through the evaluation standards. Many districts used a rotation variation such as “major, minor, minor” in which teachers were evaluated first on all standards (usually four) for the major year and then on two standards each in the following two minor years. Other districts used a type of electronic method of collecting observation information and producing the final evaluation.
With seventeen school districts in Nevada, each coterminous with a county, seventeen different systems and approaches existed for the evaluation of teachers and administrators. This diverse approach to teacher evaluation across the state led to a desire from policy makers and educators to align the evaluation system by creating a statewide system of evaluation. The new NEPF legislation required site administrators to evaluate teachers on ten standards every year: five in the area of instruction and five in the area of professional responsibilities.

A number of studies have examined performance evaluations from the teacher’s perspective but very little inquiry has been conducted into teacher evaluation from the viewpoint of the person responsible for carrying out the process: the school principal. According to Darling-Hammond (2012),

Strong evaluation systems need principals and other evaluators with deep knowledge of teaching and learning, as well as an understanding of how to evaluate teaching, how to give useful feedback, and how to plan professional development that supports teacher learning. The lack of such knowledge and training has been a major problem for the validity, fairness, and utility of many teacher evaluation systems. (p. 28)

With the new rules in place for the 2015-2016 school year, the knowledge, attitudes, assumptions, and beliefs that principals held were not well understood with regard to teacher evaluation using the NEPF. Little was known about school leaders’ approaches to the process or their beliefs in regards to evaluation as an education reform effort. The changes that school administrators were making due to the new evaluation system were
not explicit. It was not known how they were managing the transition to a statewide system or how they were using evaluation information.

**Purpose of the Study**

The purpose of this study was to develop a better understanding of the beliefs, attitudes, assumptions, and knowledge that school principals held about teacher evaluation as related to the new evaluation system, the NEPF. The goal was to understand school site leaders’ perceptions about the process of conducting evaluations. Concepts such as beliefs about evaluation, scheduling, collecting and organizing evidence, holding reflective conversations, giving feedback, making judgments, and other aspects of evaluation were examined to understand how school principals perceived their own knowledge, beliefs, and approaches to the evaluation of teachers in the new system.

**Research Questions**

This study examined the mental models of sixteen principals to uncover their knowledge, beliefs, attitudes, and assumptions regarding teacher evaluation. The principals were chosen from four non-metropolitan school districts in Nevada. The research questions that guided this research were:

1) What mental models (beliefs, assumptions, attitudes, and knowledge) do principals hold about the role of teacher evaluation as a vehicle for change in schools; and,

2) Do different school contexts (elementary, middle, and high school) influence the mental models that site administrators hold relative to teacher evaluation?

Because the organization influences the individual and the individuals’ mindsets influence the organization, the mental models of the principals were examined through
the lens of how organizations learn. Organizational learning, based on the work of Chris Argyris and David Schö'n (1996) and Peter Senge (2006), was the larger framework of the study. Edmondson and Moingeon (2004) wrote,

> We view the contributions of Senge and Argyris as complementary parts of a theory of intervention that focuses on examining and developing mental models. Our analysis of relationships among different foci in the organizational learning literature suggests that this intervention strategy offers critical leverage for reinterpreting organizational situations and changing persistent routines. (p. 33)

Mental models play a significant role in how leaders act and react in the context of their schools.

Discovering mental models and understanding beliefs, attitudes, assumptions, and knowledge of individuals is best accomplished through dialogue and reflection. Therefore, the research design chosen for this study was qualitative, using the constructivist grounded theory approach as described by Charmaz (2014). Principals were interviewed using a set of questions designed to explore their processes, approaches, thinking, and understanding of the teacher evaluation process. The interviews were analyzed for themes and commonalities as well as anomalies. A coding system was implemented to identify the themes. Through a constant comparative method, the themes were checked for consistency. Once there was repetition of the themes and no new themes were appearing, the themes were categorized. The categories then were examined for emerging theories.

**Significance of the Study**

In terms of successful teacher evaluations, school principals are key (Bradley, 2014). One of the most impactful tasks of the school leader is how teacher evaluations are
conducted. According to Cosner, Kimball, Barkowski, Carl, and Jones (2014), “it is the school principal who must lead and oversee the change process at the school level” (p. 78). Any school initiative must be monitored by the principal to ensure quality and compliance. This monitoring is often accomplished through the lens of teacher evaluation.

Understanding the mental models of school leaders regarding teacher evaluation may provide an important body of information that could unearth ideas for enhancing practice, inform future decisions about professional development, and identify other support mechanisms for school administrators. Operating from the assumption that school leaders want to be effective, conclusions from this study may be useful in providing information that enhances their effectiveness.

**Assumptions**

Certain assumptions came into play during the course of the study. It was assumed that respondents participated in the evaluation process with integrity and were honest in reporting their perceptions. It also was assumed that teacher evaluation in the schools was part of a larger system of improvement where mental models played a role in the success of the learning organization.

**Limitations**

The first limitation to this study was the method of identifying participants. The sample was created by purposeful sampling rather than by a true random sample. In addition, the fact that the researcher was known, though not personally, to many of the participants may have created limitations to the openness of the answers during the interview process. A further limitation may have been that, at the time of the interviews, principals were beginning their second year using the new state evaluation system. This
situation could have had a positive impact on responses in that administrators may have had the opportunity to make comparisons with previous procedures; however, it could also have had a negative impact in that principals could have been struggling with the new schedules and requirements of the system. Finally, the principals in this study were all selected from small- to mid-sized school districts. The information gained was limited to that context and therefore not generalizable to larger school districts.

Ethical Considerations

One consideration in this study was the issue of privacy. Because the districts included in the study were relatively small with close professional connections, every effort was made to maintain confidentiality of data and to practice neutrality of reporting. Participants were made aware that names were removed from responses during analysis and that the results of analysis were reported in the aggregate. However, quotes were used, which could increase a possibility of recognition.

Definition of Terms

The following terms will be encountered during the course of this study. The definitions addressed are meant to clarify words or phrases used in the context of this particular project as related to schools and education.

- **Assumption**: Something that is believed to be true or probably true but that is not known to be true (Merriam Webster, 2016).
- **Attitude**: A feeling or way of thinking that affects a person's behavior (Merriam Webster, 2016).
- **Belief**: A feeling of being sure that someone or something exists or that something is true (Merriam Webster, 2016).
• **Constant comparison method:** According to Charmaz (2014),

A method of analysis that generates successively more abstract concepts and theories through inductive processes of comparing data with data, data with code, code with code, code with category, category with category, and category with concept. In the last stages of analysis, researchers compare their major categories with those in relevant scholarly literatures. Comparisons then constitute each stage of analytic development. (p. 342)

• **Constructivist grounded theory:** According to Charmaz (2014),

A contemporary version of grounded theory that adopts methodological strategies such as coding, memo-writing, and theoretical sampling of the original statement of the method… Thus, constructivist grounded theorists attend to the production, quality, and use of data, research relationships, the research situation, and the subjectivity and social locations of the researcher. Constructivist grounded theorists aim for abstract understanding of studied life and view their analyses as located in time, place, and the situation of inquiry. (p. 342)

• **Deuterolearning:** According to Argyris and Schön (1996), “second-order learning, or ‘learning how to learn’…Organizational deuterolearning is critically dependent on individual deuterolearning” (p. 29).

• **Double-loop learning:** According to Argyris and Schön (1996), “learning that results in a change in the values of theory-in-use, as well as in its strategies and assumptions…refers to the two feedback loops that connect the observed effects of action with strategies and values served by strategies” (p. 21).
• Espoused theory: According to Argyris and Schön (1996), “the theory of action which is advanced to explain or justify a given pattern of activity” (p. 13); that is, the formal description of a plan or activity, as opposed to a theory-in-use.

• Holonomy: “From the Greek holos – whole - with the suffix on suggesting a part” (Koestler, 1972, p. 112). The concept that every part influences the whole and that the whole influences the part; the relationship between part and whole, “displaying both the independent properties of wholes and the dependent properties of parts” (Koestler, 1972, p. 112).

• Knowledge: Information, understanding, or skill that you get from experience or education (Merriam Webster, 2016).

• Mental models: According to Senge (2006), “Mental models are deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action” (p. 8).

• Organizational Learning (OL) versus Learning Organization (LO): In essence, organizational learning is related to the process by which “individuals in organizations take action to develop and refine their cognitive maps – for example, their ‘theories-in-use’ (Argyris & Schön, 1974) or ‘mental models’ (Senge, 1990) – and thereby become more effective decision makers” (Edmonson & Moingeon, 2004, p. 28). Whereas, the learning organization is related to the structure where “engaging individuals in reflecting upon and developing their own thinking processes is an essential component of creating learning organizations” (Edmonson & Moingeon, 2004, p. 27).
• Single-loop learning: According to Argyris and Schön (1996), “instrumental learning that changes strategies of action or assumptions underlying strategies in ways that leave the values of a theory of action unchanged” (p. 20).

• Theories-in-use: According to Argyris and Schön (1996), “the theory of action which is implicit in the performance of that pattern of activity” (p. 13); that is, the actual way in which the action is performed as opposed to the way the action is described in the espoused theory.

Summary

Chapter One provided an introduction to the study, including background, statement of the problem, research questions, assumptions, limitations, and definitions. A qualitative study was conducted of the mental models that school principals held regarding a new teacher evaluation system. Sixteen principals from small- to medium-sized districts in the State of Nevada were invited to interview. The understanding of mental models was placed within the larger context of organizational learning. A constructivist grounded theory approach was utilized during the collection and analysis of interview data. A constant comparative method was followed to code, compare, categorize, and theorize using the interview data. A final theoretical sampling informed the final write-up of results.
Chapter Two: Review of Relevant Literature

A review of literature was conducted to provide background for this research study. The focus of the research was on mental models as they relate to organizational learning and organizational change, particularly in education. In addition, review of the literature was conducted as related to school principals, school improvement, and teacher evaluation. The review of literature consists of a discussion of holonomy, an overview of organizational learning, and a discussion of mental models within that context. Current trends in teacher evaluation are explored and, finally, a rationale for the use of constructivist grounded theory as an appropriate lens for data collection and analysis is presented.

Holonomy

The nested nature of educational systems presents an important consideration for the study of principals’ mental models because of the holonomous relationship between the individual and the organization. The term “holon,” first coined by Arthur Koestler, comes from the Greek holos, meaning whole, with the suffix on referring to a part. Koestler (1972) described the concept of holons as follows:

These sub-wholes – or “holons”, as I have proposed to call them – are Janus-faced entities which display both the independent properties of wholes and the dependent properties of parts. Each holon must preserve and assert its autonomy, otherwise the organism would lose its articulation and dissolve into an amorphous mass – but at the same time the holon must remain subordinate to the demands of the (existing or evolving) whole. (p. 112)
The concept of holonomy can be extrapolated to education and the importance of the relationship that exists between the individual and the system. Costa and Kalick (1995) described the symbiotic connection between the units of the system and the individual in the following terms:

The holonomous organization, therefore, consists of two components: parts and whole. One function of the holonomous organization is to support people in becoming autonomous and self-actualizing. This implies that each unit in the organization—individuals, classrooms, schools, districts, and the state—will become self-evaluating, self-renewing, and self-modifying. Yet the concept of holonomy transcends individual autonomy and supplies a missing link between the individual and the larger organization. (p. 5)

Holonomy plays a role in understanding how the mental models of individual principals affect the school as a learning organization. Likewise, the school as an organization (a collection of individuals) influences the mental models of the school leader and other individuals within the school. Because one of the main responsibilities of a school principal is to conduct the supervision and evaluation of teachers, the mental models that principals hold about teacher evaluation affect the process. Gaining insight into the beliefs, attitudes, assumptions, and knowledge (the mental models) from which principals operate regarding the evaluation process can provide the missing link described by Costa and Kalick (1995) that exists between the individual (the principal) and the larger organization (the school).

This holonomous relationship is especially significant during the course of change. Edmondson and Moingeon (2004) concluded,
We view the contributions of Senge and Argyris as complementary parts of a theory of intervention that focuses on examining and developing mental models. Our analysis of relationships among different foci in the organizational learning literature suggests that this intervention strategy offers critical leverage for reinterpreting organizational situations and changing persistent routines. (p. 33)

As policy makers consider teacher evaluation as a reform effort to promote changes in instruction and student learning, understanding the mental models of the leaders of that change effort brings significant emphasis to the process. As indicated by Cosner et al. (2014),

In an ongoing quest to improve student learning across the US, policy makers at the federal level have embraced an educator effectiveness agenda. This new policy agenda targets the improvement of educator practice and effectiveness as a key mechanism for improving student learning. (p. 76)

As such, teacher evaluation is now considered a high stakes leverage point that has implications for improvement system-wide in education. To provide context for considering system-wide improvement in education, organizational learning research will be explored.

**Organizational Learning**

Organizational learning includes a number of variations in management literature. Two main approaches have been provided by the seminal research of Chris Argyris and David Schön (1978) and Peter Senge (1990). Argyris and Schön (1996) considered organizational learning mainly through the lens of process, encompassing feedback loops and error analysis. Senge’s (2006) approach was through the learning organization’s
structure of systems thinking and the effects of current policies. Both incorporate a “cognitive approach to intervening in organizations to improve their adaptability and effectiveness” (Edmondson & Moingeon, 2004, p.34). Argyris and Schön (1996) provided the cognitive aspect through understanding "theories-in-use versus theories of action" while Senge (2006) used the cognitive psychology term mental models. Both sought to understand the knowledge, attitudes, beliefs, and assumptions of organization members as a crucial component of organizational change and growth.

**Organizational learning: A process.** Argyris and Schön (1996) identified two philosophies of organizational learning (OL). The first branch was described as “prescriptive, practice-oriented, value committed, sometimes Messianic, and largely uncritical” (Argyris & Schön, 1996, p. xix). The second branch was identified as having the characteristics of being “a rich topic for scholars, distant from practice, skeptical of the first-branch claims, nonprescriptive, and neutral with respect to its definition of learning” (Argyris & Schön, 1996, p. xix). Both branches include certain commonalities which exert positive influences on the learning of the organization. Argyris and Shön’s approach to organizational learning was intended to cut across the two branches by examining the elements that unite them. Argyris and Schön (1996) emphasized the importance of organizational theories of action and the distinction between single- and double-loop learning. They also included the behavioral world of the organization, that is, the theories-in-use practices of individuals (also considered mental models) that reinforce the behavior which then reinforces the theory-in-use. According to Argyris and Schön (1996), “If theorists of organizational learning seek to be of use to practitioners, they must somehow link organizational learning to the practitioner’s thought and action” (p. 6).
It is this linkage of organizational learning to practitioner thought and action that is at the heart of organizational learning.

Argyris and Schön (1996) described organizational learning from an inquiry standpoint and alluded to the idea of holonomy by observing,

We recognize the complex interactions that occur between individual and organizational learning. We see the causal arrow pointing in both directions: the learning of the individuals who interact with one another is essential to organizational learning, which feeds back to influence learning at the individual level. (p. xxii)

The concept of inquiry is important in this model because the underlying assumption is that “inquiry does not become organizational unless undertaken by the individuals who function as agents of an organization to which they belong” (Argyris & Schön, 1996, p. 11).

Knowledge becomes organizational when the output of the organizational inquiry represents a change in the thinking and acting of individuals within the organization. The basic concept is that changes in thinking and acting yield changes in organizational practices which is the theory of action. A theory of action development flowchart is illustrated in Figure 1.
Figure 1. Development of a theory of action. The graphic above represents the development of a theory of action. S represents the existing situation requiring change. Study of the situation reveals the underlying assumptions that make it plausible that a certain action strategy (A) will produce the desired consequence (C). The values attributed to the consequence (C) that make it a desirable outcome are revisited in light of situation S. Adapted from Argyris and Schön (1996).

Figure 1 represents the driving forces within a theory of action as described by Argyris and Schön (1996). Situation (S) represents a set of events as it exists within the organization. Implicit in this situation are underlying assumptions, values, forces, and an existing model that make it plausible that an action strategy (A) will produce a particular consequence (C). The consequence C is an intended outcome, laden with values that make it desirable. Within the structure of the organization, to produce the expected consequence C, then A must be accomplished. In turn, the thinking and actions, or the mental models, of individual practitioners within the organization are aligned with the theory of action. However, for growth and change to occur, stakeholders need to question or to conduct inquiry in relation to the theory of action.

In the context of theory of action, Argyris and Schön (1996) identified two different forms. One form is the espoused theory which refers to the theory used to explain or justify the public theory of action; that is, the given pattern(s) of activity in the
espoused theory of action or, simply, the formal plan. An educational example might be a written procedure for collecting weekly teacher lesson plans.

The other form is referred to as the theory-in-use which describes the theory of action that is implicit in the performance of the pattern of activity in the espoused theory; that is, the actual way in which the action is performed as opposed to the way the action is described in the espoused theory. Theory-in-use represents the belief system, or mental model, used by individuals or a collection of individuals to enact the espoused theory of action (Argyris & Schön, 1996). In other words, the written or espoused plan might say one thing but individuals do not do it that way. The procedure for collecting weekly lesson plans named above might in actuality be the principal randomly viewing lesson plans displayed in classrooms as he or she walks the school.

There can easily be a mismatch between the espoused theory of action and the theory-in-use. A further educational example of an espoused theory is Response to Intervention (RTI) as described in the performance plan of a hypothetical elementary school. This plan might describe how students who are underperforming in math would be assessed bi-monthly and placed in flexible groups to provide intervention lessons at their level. The system is described as responsive, fast acting, and individualized.

Upon investigation of RTI in practice, the theory-in-use might reveal a different pattern of behaviors. Instead of following the described process of the espoused theory of action, students are assessed less frequently, students remain in intervention groups for long periods, and individualization seldom occurs. In other words, intervention groups do not demonstrate flexibility but remain static over time, thereby defeating the purpose of the espoused theory of action. Thus, from the view of Argyris and Schön, the comparison
of these two forms of theories establishes a basis for analysis and evaluation of the theory of action. The results of the analysis are used as feedback to promote change.

**Single- and Double-Loop Learning.** The process of analyzing the theory of action that can lead to organizational learning can be accomplished through the concepts of single-loop and double-loop learning. Single-loop learning is associated with one chain of feedback to key stakeholders. Single-loop learning occurs when feedback causes changes in strategies or actions that are standard operating procedures within the organization but does not change the values and norms of the organization. Single-loop learning is typical of incremental improvement in the system. For example, school personnel may acquire information regarding safety issues in the pick-up and drop-off routines at an elementary school. Based on this information, school personnel investigate the problems and change the routines accordingly. Monitoring of the new routines indicates that the solution was the correct one, thus solving the problem. The action taken corrected the detected issue in keeping with existing norms and values of the organization; that is, the underlying values and norms were not changed.

Figure 2 illustrates the cycle of single-loop learning where a correction is made through observation, assessment of possible corrections, design of alternatives, and implementation of action strategies. There is no examination of organizational values and norms. There is not a feedback loop to inform the organization of any possible implications for adapting changes to fundamental beliefs or assumptions. There is only a single feedback loop.
Figure 2. Single-Loop Learning. The graphic above demonstrates the process of single-loop learning within an organization. Single-loop learning is characterized by an organization, or individuals within, using information to make decisions, taking action, and then checking to see that the action is working, without making changes to the fundamental organizational values, beliefs, assumptions, or norms. There is only one feedback loop. Adapted from Argyris and Schön (1996).

Double-loop learning occurs when organizational learning results in change in the actual values, assumptions, and strategies of the theory-in-use of the organization. Double-loop learning is represented by two feedback loops: one loop brings information and learning to the organization, causing changes in strategies and actions through confirmation of the results of the action taken. The other loop connects the results of the strategies and actions back to the organization to provide information and learning that fundamentally change the values and assumptions of the organization.

As an example, let us return to an elementary school that uses a system of professional learning teams to examine student learning in math. Each team examines
data related to its students, develops plans for corrective actions, and implements the plan. This is the first loop. The extra intervention opportunities seem to be increasing the learning of the students in one team but not in another. School leadership then provides opportunities for each team to share its data and strategies for increasing student learning; this is the second loop. The teams engage in discussions and information sharing. The teams discuss the teaching strategies that worked and decide as a school to shift the focus of math instruction to include more student conversation and problem solving. Over time, data indicate that math understanding is increasing among all students. As a result, teachers and administrators shift their fundamental expectations of the teaching of math to include a significant amount of student discourse. The school community as a whole has now shifted its thinking, its values, and beliefs. The school community’s theory-in-use, that is, its mental model, has changed.

In this example, the double-loop structure allowed the opportunity for feedback to provide information that promoted organizational learning. This learning was associated with change in the fundamental beliefs and expectations of the school staff as a whole. That is, the mental model related to quality math instruction changed. Figure 3 provides a visual of the processes of single- and double-loop learning within an organization. The cycle on the left represents single-loop learning which is characterized by an organization, or individuals within, using information to make decisions, taking action, and then checking to see that the action is working. The visual extends to the right to represent how double-loop learning contains the added dimension of collective assessment and examination of the value of the actions from the single-loop. Reflection may lead to changes in the collective values, beliefs, norms, and assumptions underlying the theories-
in-use, or mental models, of individuals or whole groups. Indeed, without double-loop feedback, the theory-in-use cannot change (Argyris & Schön, 1996).

**Figure 3.** Single- and Double-Loop Learning. The graphic above demonstrates the processes of single- and double-loop learning within an organization. Single-loop learning is characterized by an organization, or individuals within, using information to make decisions, taking action, and then checking to see that the action is working. Double-loop learning contains the added dimension of collective assessment and examination of the value of the actions from the single-loop. Reflection may lead to changes in the collective values, beliefs, norms, and assumptions underlying the theories-in-use, or mental models, of individuals or whole groups. Adapted from Argyris and Schön (1996).

Double-loop learning is associated with the parallel possibility for identification of first and second order errors. First order errors are related to the theory-in-use process that can be identified and corrected. That is, first order changes are associated with incremental improvement within the system. For example, a bookkeeping error is detected and can be corrected.
Second order errors are identified as the result of deep questioning, information gathering, and reflection during double-loop learning. Conscientious comparison of the espoused theory with the theory-in-use can reveal gaps. When double-loop learning inquiry is successful, it provides important and valid information about second order errors and promotes valuable learning that may lead to a change in fundamental values and beliefs. When double-loop learning is not successful, a second order error may occur which perpetuates a systemic problem because underlying values and beliefs are not questioned.

To return to the elementary school example, when the learning teams honestly examined their expectations and practices for math instruction, they identified the gap between their theory-in-use and the espoused theory. They conducted an inquiry related to norms and values regarding math instruction and ended up making significant changes for the better. Second order errors were detected and corrected. Realizations such as these that cause positive changes are linked to organizational learning and the process of how an organization learns.

*Deuterolearning.* How an organization learns is part of a learning system and an additional consideration in the organizational learning process. Argyris and Shôn (1996) borrowed the phrase deuterolearning from Gregory Bateson’s 1972 anthropological and psychological study of mammal communication where patterns of how mammals learn to learn are described and analyzed (Bateson, 1972). Argyris and Shôn (1996) applied the term deuterolearning to the phenomenon of organizational learning where individuals and groups within an organization learn how they learn together. They consciously effect
change in their beliefs and practices, or their mental models, with the goal of eventually bringing change to an entire organization.

According to Argyris and Schön (1996), an organizational learning system is based on both structural and behavioral features that provide a coherent environment in which to conduct organizational inquiry. Structural features of the organization include channels of communication, information systems, the physical environment, procedures and routines, and systems of incentives. Behavioral features include the “qualities, meanings, and feelings that habitually condition patterns of interaction among individuals within the organization in such a way as to affect organizational inquiry” (Argyris & Schön, 1996, p. 29). In other words, does the organization maintain a collaborative, supportive, and inquisitive environment and culture of communication that enhances collective inquiry? How an organization learns how to learn is representative of, indeed contingent upon, the holonomous interaction between the collective understanding of these procedural and behavioral features and the way individuals in an organization understand these features. The mental models maintained by individuals play a large part in the deuterolearning process of an organization.

In contrast to Argyris and Schön, where mental models are implicitly contained in the organizational learning process through theories-in-use, Senge (2006) explicitly identified mental models as one of the five disciplines to which organizations must attend to become a learning organization. The meaning of discipline in the context of Senge’s view of organizational learning comes from Latin (*disciplina*, to learn) and refers to “a body of theory and technique that must be put into practice” (p. 10). However, Senge considered these five disciplines to be personal disciplines as opposed to management disciplines in that they focus on how people interact and learn from one another. The five disciplines are: Personal Mastery, Mental Models, Building Shared Vision, Team Learning, and Systems Thinking.

The five disciplines are interrelated and continually influence each other. As one discipline changes, there will most certainly be a consequence that can be intended or unintended and that can be positive or negative in another area of the system. Therefore, it is crucial to continuously consider all the disciplines and the patterns of interrelationships among the key elements of the system. Each discipline will be discussed individually, though in practice “it is vital that the five disciplines develop as an ensemble” (Senge, 2006, p. 11). Figure 4 represents the five disciplines with systems thinking as the backdrop for the other disciplines.
Personal mastery. Senge (2006) made it clear that developing personal mastery in the context of organizational learning is not to gain dominance over people or things but to continually seek a high level of proficiency within a certain area of focus by becoming committed to lifelong learning. Senge (2006) stated,

“Personal mastery” is the phrase we use for the discipline of personal growth and learning. People with high levels of personal mastery are continually expanding...
their ability to create the results in life they truly seek. From their quest for continual learning comes the spirit of the learning organization. (p.131)

The central principle of the discipline of personal mastery revolves around the idea of *creative tension*. Creative tension includes two underlying concepts: one, continually clarifying and deepening a personal vision; that is, each person must individually identify what is most important in life. Second, continually learning how to see current reality more clearly; that is, seeking a true understanding of surrounding situations as well as one’s personal contribution to them. The push and pull created between what is important and what is reality sets up the condition of creative tension.

Developing the ability to balance creative tension in life, including managing emotional tension and structural conflict, represents the essence of personal mastery and changes the way one views failure. Senge (2006) defined failure in this context as “a shortfall, evidence of the gap between vision and current reality. Failure is an opportunity for learning” (p. 141). Personal mastery is learning that goes beyond knowing information; it requires “lifelong generative learning. And learning organizations are not possible unless they have people at every level who practice it” (Senge, 2006, p. 132). As a leader, supporting the development of personal mastery in individuals leads to a higher level of happiness in each person through self-fulfillment along with commitment, initiative, responsibility, quicker learning, and subsequently the confidence to support others in their learning: All desirable characteristics in a learning organization.

The intersection of personal mastery and the fifth discipline, systems thinking, brings other aspects of personal mastery to light. According to Senge (2006), people with high levels of personal mastery are able to integrate reasoning and intuition; to
continually see more of a connectedness to the world; to employ compassion; and to commit to the whole, to something with a larger purpose. Senge (2006) cautioned against requiring personal mastery programs but to instead foster a culture of personal mastery by working hard to create a climate where it is safe for all people to create visions, to practice inquiry and commitment to truth (understanding reality), and to challenge the status quo. A leader can promote this climate by being a model, by practicing the principles of personal mastery and looking for them within the organization.

**Mental models.** Senge (2006) defined mental models as “deeply engrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action” (p. 8). Mental models may unconsciously influence the way an entire organization perceives the outside world, how decisions are made, and how the organization acts upon those decisions, collectively or individually. Through working with mental models, an organization can make changes by unearthing underlying beliefs and assumptions and then examining them via inquiry and “learningful” conversations (Senge, 2006, p. 8). Mental models have a powerful influence on what is done because they affect what is seen and believed. For example, a teacher who maintains a mental model of “girls are not good at math” may very likely treat girls differently than boys in the classroom. The teacher may call on boys more often, give girls less challenging problems to solve, or overlook girls for awards and recognitions.

Often, mental models are hidden, existing below the surface of an individual’s or a group’s consciousness. If these mental models are not identified and examined, they remain unchanged, which can eventually exacerbate problems or lead to counterproductive decisions. However, according to Senge (2006), “if mental models can
impede learning...why can’t they also help accelerate learning? This simple question became, over time, the impetus for the discipline of bringing mental models to the surface and challenging them so they can be improved” (p. 167). Thus, integrating the discipline of mental models consistently with systems thinking allows the LO to expose existing assumptions about current professional practices, to examine and reflect on them, and to consider how they might be updated, improved, and operationalized.

Building shared vision. According to Senge (2006), “the practice of shared vision involves the skill of unearthing shared ‘pictures of the future’ that foster genuine commitment and enrollment rather than compliance” (p. 9). A shared vision of the future in an organization creates the desire to learn and grow because of a mutual interest in a lofty goal. People learn because they have both a stake and an interest in contributing to the shared vision. Senge (2006) suggested that, simply stated, a shared vision is the answer to the question, “What do we want to create?”

A shared vision takes an organization deeper than just having an idea. It “creates a sense of commonality that permeates the organization and gives coherence to diverse activities” (Senge, 2006, p. 192). Individuals may create personal visions, but a shared vision becomes powerful when each individual is committed to the same vision and expects each other to maintain that vision. Senge (2006) indicated that people want to be connected to a higher purpose or important undertaking; therefore, they become invested through the process of building shared visions.

A shared vision within the context of systems thinking is crucial because it provides a picture for what members of an organization want to create. Integrating shared vision with systems thinking reveals how the current reality was created in the first place.
Understanding how and why the current vision exists gives possibility to not only understanding both past issues and successes, but how to move forward with next steps by re-envisioning the future.

**Team learning.** According to Senge (2006), “team learning is the process of aligning and developing the capacity of a team to create the results its members truly desire” (p. 218). A crucial component of team learning is dialogue. Dialogue goes beyond any one person’s understanding and refers to the act of individuals being able to suspend their assumptions in order to think together through conversation. Dialogue is different from discussion. Discussion is more akin to a ping-pong game, where ideas are jettisoned back and forth, analyzed and dissected. This type of group conversation can be useful and does have its place, but, as with a ping-pong game, the purpose of discussion is to win a point, to have a certain view be accepted by the group. In contrast, dialogue allows individuals to gain insight, expand their levels of understanding, or explore avenues of thought stimulated by group conversation that would not be accessible to a person alone. Groups who are able to accomplish this type of exchange develop a free-flow of ideas that allow team learning to increase. It is a metacognitive activity: “In dialogue people become observers of their own thinking” (Senge, 2006, p. 224).

Becoming more sophisticated in the process of dialogue can lead to the ability of a team to identify incoherent thinking as well as productive thinking. Working together through the incoherence of thought brings more and more coherence and deeper understanding, leading to team learning.

According to Senge (2006), discussion and dialogue are necessary skills for team learning. In discussion, viewpoints are brought forward and defended and decisions are
made. Dialogue is a vehicle whereby different viewpoints are brought forward with the intention of discovering a new viewpoint through collective reflection and inquiry. In a learning organization, it is necessary to distinguish between these two approaches and to become adept at engaging in the appropriate function for the appropriate purpose. Team learning is a group skill and requires focused practice to be effective. As Argyris and Schön (1996) also discussed in deuterolearning, a team must together learn how to learn.

Senge (2006) pointed out that “the tools of systems thinking are also important because virtually all the tasks of management teams--developing strategy, shaping visions, designing policy and organizational structures--involve wrestling with enormous complexity” (p. 249). It is through dialogue (and perhaps discussion) that teams learn the complex language of systems thinking and practice addressing all the disciplines at the same time. Each person brings his or her mental models to the table and grapples with how to deal with complex issues in a more objective and dispassionate way.

**Systems thinking.** Senge (2006) referred to systems thinking as the fifth discipline because it is the “conceptual cornerstone that underlies all of the five learning disciplines” (p. 69). It is a discipline for “seeing wholes” (Senge, 2006, p. 68), made up of a set of general laws and principles, tools, and techniques for understanding systems.

Senge (2006) described eleven laws of systems thinking, synthesized from many systems thinking theorists. The laws are described briefly as follows:

1) Today’s problems come from yesterday’s solutions. Ineffective solutions of the past can merely be shifting a problem from one part of the system to another, causing the problem to go undetected.
2) The harder you push, the harder the system pushes back. This law is derived from *compensating feedback*. Overcompensating for a problem can cause problems elsewhere in the system.

3) Behavior grows better before it grows worse. Short term solutions may give way to long-term problems.

4) The way out usually leads back in. Solutions that are too familiar may not be the right ones for the issue.

5) The cure can be worse than the disease. A solution taken in haste or familiarity can develop a dependency on outside resources, shifting the burden of the problem and possibly perpetuating it.

6) Faster is slower. Moving too fast can cause a system to bog down or even collapse.

7) Cause and effect are not closely related in time and space. It may be necessary to look more deeply for causes to a problem in a complex system.

8) Small changes can produce big results – but the areas of highest leverage are often the least obvious. Consider the forces at play in the system to identify changes that can make big differences in the long run.

9) You *can* have your cake and eat it too – but not at once. More than one dilemma can be improved over time, focusing on one at a time.

10) Dividing an elephant in half does not produce two small elephants. Consider issues from different aspects – those that are closest to the problem as well as from an overall perspective.
11) There is no blame. No one entity, inside or out of the system, is a culprit. A system survives on inter-relationships.

Senge (2006) described systems thinking based on seeing “patterns of change rather than static snapshots” (p. 68). Senge encouraged identifying the difference between detail complexity, that is, identifying many short-term variables, and dynamic complexity. According to Senge (2006), dynamic complexity deals with understanding more subtle, long-term effects by “seeing interrelationships rather than linear cause-effect chains and seeing processes of change rather than snapshots” (p. 73).

The tools to understanding systems thinking lie in the concepts of feedback and delays. According to Senge (2006), feedback is “any reciprocal flow of influence” (p. 74). In this definition, feedback is seen as a continual circle of cause and effect as opposed to a linear perspective. Senge identified two major types of feedback: reinforcing feedback, which amplifies growth; and balancing feedback, which can slow a system down as needed. Delays are described as “interruptions between your actions and their consequences” (Senge, 2006, p. 88). Delays can be positive or negative, depending on how they are detected and managed.

Techniques of systems thinking include combining several principles. Systems thinking entails organizing detail complexity to understand dynamic complexity, identifying systems archetypes to see existing structures, applying the tools of feedback and delays, and maintaining a shift from only linear, short-term thinking to include long-term, broader perspectives.
Systems thinking requires the integration of the other four disciplines. The interconnectedness of all of the disciplines is what leads to developing a learning organization.

**Seven organizational learning disabilities.** Even with putting the five disciplines in place, there is no guarantee that an organization will automatically succeed. Senge (2006) presented seven learning disabilities that exist to some degree in all learning organizations and which, if left unidentified and unaddressed, can undermine and prevent successful learning. Senge’s organizational learning disabilities are briefly described below:

1) **I am my position:** Individuals confuse their personal identities with their positions. This confusion leads to people not being able to see beyond the boundaries of their position, limiting understanding of the big picture or greater purpose, and therefore not feeling a sense of responsibility for results when all positions work together.

2) **The enemy is out there:** A by-product of *I am my position* where an external agent is blamed for things that go wrong. The external agent can be something or someone within an organization or without.

3) **The illusion of taking charge:** Being proactive is a popular notion. However, it is common to confuse proactiveness with reactiveness. Aggressively taking action against the enemy out there is reactive. Proactiveness is a product of thinking, not reacting to an emotional state. Reactively taking charge at the wrong time can exacerbate issues.
4) The fixation on events: Seeing only events and reacting to them without considering the longer-term patterns of change leading up to them can prevent understanding the underlying causes. This leads to the risk of repeating negative events or failing to identify situations that develop gradually over time.

5) The parable of the boiled frog: Learning to see evidence of slow, gradual processes that may be detrimental in the end. Placing a frog in boiling water will cause it to try to escape. Placing a frog in room temperature water and heating it slowly will cause the frog to remain until it is too late. The frog’s internal apparatus is not geared to detecting the slow changes. A learning organization must be attuned to gradual changes and what they mean lest it suffer the fate of the boiled frog.

6) The delusion of learning from experience: Directly learning from trial and error can be limiting and sometimes impossible because the results are not always readily accessible. Being able to anticipate and imagine the long-term effects of decisions is therefore crucial.

7) The myth of the management team: Management teams do not function the way people assume they should. They tend to squelch disagreement, lay blame, and break down under pressure. They may be useful to manage basic issues but do not do well with complicated or difficult situations.

In 2006, Senge indicated that “the discipline of managing mental models - surfacing, testing, and improving our internal pictures of how the world works - promises to be a major breakthrough for building learning organizations” (p. 163). The
understanding of mental models may well be an antidote for the seven disabilities. The next section will look more closely at mental models in light of recent research.

**Understanding Mental Models**

The concept of mental models has its roots in philosophy and cognitive psychology and examples can be tracked back to the nineteenth century (Johnson-Laird, 2004). Pepper (1942) explored the concept of world hypotheses, based on root metaphors, to enable a better understanding and interpretation of the world. Through these world hypotheses, Pepper (1942) laid a foundation for how the world can be understood through models as “they purport to inform us about the structure of the world” (p. 74). Craik (1943) described how people create internal mental models derived from external reality to navigate the world around them and react in emergency situations. More recently, Preskill and Torres (1999) considered mental models to be what guide people in their everyday lives.

In order to understand how to support learning and change, it is important to understand peoples’ mental models. This understanding provides insight into how people relate to their work, to others around them, and why they make the decisions they do. Preskill and Torres (1999) explained how important the understanding of mental models is to the learning organization:

The exploration of individuals' mental models, knowledge structures, cognitive maps, schema, frameworks, and paradigms helps us understand the role of memory in learning and how future action may be predicted. Without examining what underlies our thinking, we are prone to continue operating in old ways, limiting the potential for learning and change. (p. 66)
Argyris and Shõn (1978) and Senge (1990) established the importance of mental models in the development of the learning organization and its leadership. Senge (2006) stated, “Until the gap between my espoused theory and my current behavior is recognized, no learning can occur” (p. 177). The study of mental models is the mechanism for filling the gap between espoused theory and behavior, leading to learning.

Rook (2013) offered the following definition of the concept of an individual mental model: “A concentrated, personally constructed, internal conception, of external phenomena (historical, existing or projected), or experience, that affects how a person acts” (p. 42). Mental models affect how a leader understands current initiatives, how a leader perceives the organization’s understanding of initiatives, and how a leader acts on perception. An effective leader understands mental models.

Leithwood et al. (2004) described how important effective school leadership is: Of all the factors that contribute to what students learn at school, present evidence led us to the conclusion that leadership is second in strength only to classroom instruction. Furthermore, effective leadership has the greatest impact in those circumstances (e.g., schools “in trouble”) in which it is most needed. This evidence supports the present widespread interest in improving leadership as a key to the successful implementation of large-scale reforms. (p. 69)

As Leithwood et al. (2004) indicated, the importance of school leadership is crucial, but to exert positive impact on a school, leadership must be effective. Understanding mental models is an important aspect of being an effective leader. As Senge and Sterman (1991) pointed out, “many argue that improving the mental models of managers is the fundamental task of strategic management” (p. 1007), but that “flaws in mental models
cannot be corrected until mental models become more explicit” (p. 1010). The following studies reinforce the theory that transparency of mental models plays a significant role in the success of improvement efforts in schools.

In 2005, Ruff and Shoho examined the mental models of three school principals as related to instructional leadership. The three principals included one novice principal, a principal with 6 years of experience, and a principal with 20 years experience who was a Nationally Distinguished Principal as well as a Texas Association of Elementary School Principals Principal of the Year. All three principals led successful urban schools. Interviews, observations, and review of relevant artifacts were used to collect information regarding the mental models of the principals. Ruff and Shoho’s (2005) conclusions described a common cognitive structure in which the “issues, conditions, routines, and words used in describing all three principals were similar” (p. 571). However, even though there was similar understanding of routine functions amongst the principals, each principal acted differently based on his or her mental model. Where the experienced principal had a more developed sense of her own mental model of leadership, actions resulted in the increased success of the school. Results showed the usefulness of the concept of mental models to describe differences in instructional leadership; “because of their capacity to uncover tacit assumptions, they serve to clarify abstract meanings rather than bury them further in layers of undiscussability” (Ruff & Shoho, 2005, p. 575).

Chrispeels, Burke, Johnson, and Daly (2008) examined the mental models held by central office and school leadership teams (SLTs) about team tasks related to school improvement in a Canadian school district. According to Chrispeels et al. (2008), “Negotiating potentially competing beliefs of the central office and teachers through the
narrow linkage of the principal often leads to limited understanding and lack of coherence in how to achieve goals” (p.733). The study focused for a year on tracking the progress of creating shared mental models among school leadership teams and central office goals. The study suggested that where there was a convergence in mental models; coherence developed among school leadership team members, the principal, and central office. Where there was coherence, the organizational change initiatives were moving forward at a steady pace and with a positive climate. Where there was divergence between the mental models of the leadership teams and central office, the schools and the district were struggling to meet their goals. According to Chrispeels et al. (2008),

Significantly, this study shows that with professional development, time, and safe space to surface assumptions that enable SLTs and the central office to align mental models, SLTs could also bridge previously unchartered waters between the central office and the schools in ways that can enhance organizational effectiveness, coherence, and goal attainment. (p. 747)

The sharing of leaders’ mental models in a safe environment can precipitate actions that make a difference in school change efforts.

Through a longitudinal research study, Hannay and Earl (2012) examined mental models of educators in a school district in Ontario, Canada. The district leaders implemented significant educational reforms which required adaptations of individual and collective mental models of professional practice. The study began in 2000 when four smaller school districts were required to merge into one larger district. Though the study focused on what a district could do to facilitate individual and organizational change, Hannay and Earl (2012) noted that “Without individuals reconstructing their
mental models and personal practical knowledge, external reforms will remain superficially implemented” (p. 314). The concept of holonomy was particularly important in this case: individual teachers studied their data, met with colleagues to debrief, and used the collective information to inform their practice. Principals also met in groups once a month and developed a common format to study EQAO data (the Ontario provincial testing administered in grades 3, 6, and 9). There was an identified linkage between evidence and action. The researchers concluded that through professional dialogue, the organization increased its knowledge and participants developed new mental models by reconstructing professional knowledge together.

Keskinkılıç Kara and Ertürk (2015) examined the mental models held by 61 Turkish school principals related to leadership. These principals were all enrolled in a post-graduate leadership program and had received previous leadership training. Using a data collection tool developed by the researchers, the principals were asked four open-ended questions regarding the notion of leader and were asked to draw a picture that represented their concept of leader. The four questions were:

1) How do school principals describe leadership?
2) What are the characteristics of a leader from the standpoint of school principals?
3) Do school principals believe that leadership can be acquired later?
4) How could a leader contribute to a school from the standpoint of school principals?

The analysis of the answers and the drawings indicated that the mental models did not “focus on a single point and that they consist of different models” (Keskinkılıç Kara &
Ertürk, 2015, p. 2150). The findings indicated that, even though all principals had received similar leadership training, they exhibited diverse concepts of leadership. Implications were that principals would not have the same level of efficacy in the field because they did not share the same mental models of leadership. The authors recommended that, in light of mental model theory, “Training programs conducted for educating school principals should include educational events to form accurate mental models. In addition, effort should be made to ensure that those mental models are adapted as common models” (Keskinkılıç Kara & Ertürk, 2015, p. 2151).

Review of the preceding studies reveals that where mental models were transparent, congruence of thought and action occurred. However, the goal is not to achieve a forced congruence of thought. Progress occurs through making mental models visible because it requires dialogue and understanding. According to Senge (2006),

“It’s important to note that the goal in practicing the discipline of mental models is not necessarily agreement or convergence. Many mental models can exist at once. Some models may disagree. All of them need to be tested against situations that come up. (p. 188)

The implementation of mental model theory is not simply providing an imposed, pre-determined model. The power of mental models lies in making them transparent through inquiry and the reflective process, testing them against situations that come up. It is this testing and examination of mental models that results in organizational learning (Argyris & Schön, 1996; Doyle & Ford, 1998; Ellis, Margalit, & Segev, 2012; Senge, 2006; Senge, Kleiner, Roberts, Ross, & Smith, 1994). Lozano (2011) stated “Of the myriad of human abilities, abstract thinking and creativity may be two of the most powerful and useful
ways in which humans overcome ignorance and regain awareness by generating and institutionalizing new mental models for sustainability” (p. 206).

Possibly the most significant and challenging task performed by the site principal, and one in which she should be conscious of mental models, is the process of teacher evaluation. The next section will discuss recent events leading up to the current status of teacher evaluation in Nevada.

**Teacher Evaluation as an Education Reform Initiative**

The 2001 reauthorization of the Elementary and Secondary Education Act of 1965, also known as the No Child Left Behind (NCLB) Act focused on student achievement accountability. No Child Left Behind required states to develop student achievement assessment systems for grades 3 – 8 using static proficiency levels based on adequate yearly progress. The disaggregated data from assessments were used to determine school improvement status for schools and districts. The advent of NCLB increased the focus on accountability not only for students but for teachers as well. Federal highly qualified standards for teachers raised the expectations for teacher certification. However, a highly qualified teacher is not necessarily the same as a highly effective teacher (Borman & Kimball, 2005; Tucker & Stronge, 2005). Identifying instruments that allowed administrators to accurately evaluate teachers for effectiveness in the classroom became a concern. Danielson and McGreal (2000) identified pervasive problems with most existing teacher evaluation systems, describing six main areas of deficiency:

1. Outdated, limited, evaluative criteria,
2. Few shared values and assumptions about good teaching,
3. Lack of precision in evaluating performance,
4. Hierarchical, one-way communication,
5. No difference between novice and experienced practitioners, and

Danielson and McGreal (2000) examined school districts across the United States and identified exemplary evaluation systems. They developed four recommendations based on their research to address the deficiencies listed above:

1. New evaluation systems should be directly linked to the mission of the school district,
2. New evaluation and professional development systems should be viewed as continuing processes,
3. New evaluation systems should emphasize student outcomes, and
4. There must be a commitment to allocating adequate resources to allow new systems to be successful.

In 2007, there was still a lack of knowledge regarding teacher evaluation practices. In a Regional Educational Laboratory report, Brandt, Mathers, Oliva, Brown-Sims, and Hess (2007) described their findings on studies of evaluation policies as “the few that exist are usually descriptive, outdated, and leave many questions unanswered” (p. iii).

More recently, considerable attention has been paid to the importance of teacher evaluation in school reform efforts and the principal’s role in conducting evaluations (Daly & Kim, 2010; Jerald & Van Hook, 2011; Wahlstrom, Louis, Leithwood, & Anderson, 2010).
It is widely accepted that the classroom teacher is the most influential force on student learning in the school, seconded only by the school leader (Daly & Kim, 2010; Leithwood et al., 2004; Leithwood & Riehl, 2003; Mathers et al., 2008). The evaluation of teachers has become an accepted school reform component. As Mathers, Oliva, and Laine stated in 2008, “The role of teacher evaluations has surfaced only recently as an underutilized resource that might hold promise as a tool to promote teacher professional growth and measure teacher effectiveness in the classroom” (p. 1). Since Mathers et al.’s observation, teacher evaluation has indeed become the focus of many policies.

According to the National Council on Teacher Quality (NCTQ) State of the States report from November, 2015, teacher evaluation procedures are driven by state policies. Revisions and adoptions of teacher evaluation policies have increased since 2009. The NCTQ (2015) report identifies the main findings as follow:

- 27 states require annual evaluations for all teachers, compared to just 15 states in 2009.
- 45 states require annual evaluations for all new, probationary teachers; 15 of those states specifically require that probationary teachers are observed in the classroom early in the school year.
- 17 states include student growth as the preponderant criterion in teacher evaluations, up from only four states in 2009. In 2013, 19 states had such a requirement. An additional 18 states include growth measures as a ‘significant’ criterion in teacher evaluations. Ten of those states explicitly define what significant means for the purposes of including student achievement in teacher evaluations.
- 23 states require that evidence of teacher performance be used in tenure decisions. No state had such a policy in 2009.

- 19 states require that teacher performance is considered in reduction in force decisions.

- The majority of states (28) now articulate that ineffectiveness is grounds for teacher dismissal. (p. 6)

There is evidence that educator effectiveness has emerged as a significant initiative across the United States. This focus on educator effectiveness, in the form of teacher and principal evaluation, is largely due to the significant body of research conducted regarding the importance of the teacher in the classroom and the role of the principal in assuring effective teaching (Daly & Kim, 2010; Leithwood et al., 2004; Leithwood & Riehl, 2003; Mathers et al., 2008). In addition, federal education laws of the past fifteen years have both mirrored and invited research regarding teacher evaluation. As it became clear that NCLB was not achieving desired increases in student achievement, the United States Department of Education (USDOE) developed a $4.35 billion dollar grant called Race to the Top (RTTT). Funding from this grant became available to states through application in 2009. The RTTT funding focused on four criteria:

1) Adopting standards and assessments that prepare students for success in college and the workplace;

2) Recruiting, rewarding and retaining effective teachers and principals;

3) Building data systems that measure student success and inform teachers and principals how they can improve their practices; and
4) Turning around the lowest-performing schools (USDOE, RTTT Application Workshop power point, 2009, slide 7).

Additionally, in 2012, the USDOE, under President Barack Obama and Secretary of Education Arne Duncan, developed a waiver system to allow states to expand their policies until NCLB could be revised and reauthorized by the United States Congress. According to the original waiver application of 2012:

This waiver is intended to build on and support the significant State and local reform efforts already underway in critical areas such as transitioning to college- and career-ready standards and assessments; developing systems of differentiated recognition, accountability, and support; and evaluating and supporting teacher and principal effectiveness [emphasis added]. (p. iii)

In an effort to secure funding for targeted initiatives, many states developed the teacher and site administrator evaluations needed to meet the requirements of the RTTT grant and/or the federal flexibility waiver.

In December, 2015, the revision and reauthorization of the Elementary and Secondary Education Act was completed by Congress and reissued as the Every Student Succeeds Act (ESSA). In essence, ESSA returned more control to the states in determining accountability measures, evaluation expectations, and the highly qualified (effectiveness) status of teachers and school administrators, among a plethora of other changes (USDOE, 2016). The full repercussions of the ESSA reauthorization are yet to be understood at the time of this writing.

Teacher Evaluation in Nevada. In Nevada, the federal RTTT application spurred the advent of the Nevada Blue Ribbon Task Force, established in 2010 by the
sitting Governor, Jim Gibbons (TLC White Paper, 2011). The Task Force was charged with developing a comprehensive reform plan for Nevada’s education system. This reform plan was the foundation for the Nevada State RTTT application. Though the Nevada State RTTT application was denied twice (USDOE, 2015b & 2015c), the plans put in place for the grant application became the basis for the Nevada Flexibility Waiver, approved in 2012.

As a parallel policy, Nevada Assembly Bill 222 was crafted at approximately the same time as the Nevada Flexibility Waiver. Assembly Bill 222 was passed during the 2011 Nevada Legislative Session and authorized the creation of the Teachers and Leaders Council (TLC). The TLC was required by AB 222, in alignment with the requirements of the Flexibility Waiver, to develop a statewide evaluation system for teachers and school administrators, to be approved by the Nevada State Board of Education. The new law required the evaluation system to rate educators on an effectiveness continuum: highly effective, effective, minimally effective, and ineffective. Student achievement data from statewide assessments were included as a determining factor in the rating. The TLC developed the standards and indicators with support from, among others, Dr. Margaret Heritage and the National Center for Research on Evaluation, Standards, and Student Testing (TLC White Paper, 2013). The Nevada State Board of Education approved the evaluation standards and indicators for teachers and school administrators put forth by the TLC in January, 2013. The new evaluation system was called the Nevada Educator Performance Framework, or the NEPF. The NEPF was passed into law in the Nevada Legislative Session of 2013 and focused on four goals:

- Foster student learning and growth
• Improve educator’s instructional practice
• Inform human capital decisions based on a professional growth system
• Engage stakeholders in the continuous improvement and monitoring of a professional growth system (TLC White Paper, 2013)

As of this writing, the TLC continues to function as an advisory committee for decisions related to the NEPF.

Prior to the development of the NEPF, teacher evaluation in Nevada was not unlike the rest of the United States. According to Iwanicki (2001),

In the 1980s and 1990s, a number of states and school districts developed teacher appraisal systems that analyzed teaching on the basis of what the literature defined as accepted teaching practices. These systems helped teachers and administrators develop a good understanding of accepted teaching practices. But now, teacher evaluation needs to move beyond analyzing teaching and focus on student learning. (p. 57)

Teacher evaluation was traditionally a top-down procedure based on a system that tended to keep principals out of classrooms rather than in them (Marshall, 2005). Teacher unions and district agreements established routines that often relied on a planned pre-conference, a 45 to 60 minute planned observation (more often for beginning teachers), and a post-observation conference where the administrator informed the teacher of his or her judgments.

Before the NEPF, Nevada followed a similar system. Probationary teachers were evaluated three times a year and post-probationary teachers at least once per year (Nevada Revised Statutes 391.3125, 2010). The evaluation systems were based on the
models of Robert Marzano (2011) or Charlotte Danielson (1996, 2007). The principal who followed the procedure spent hours observing and writing up the observations and providing feedback. According to Marshall (2005), “Ironically, this reduces the amount of time the saints spend in classrooms doing low-key supervision – coaching, encouraging, and gentle correction” (p. 731). Both teachers and administrators expressed frustration with this system and its lack of connection with real issues related to teaching and student learning. Principals acknowledged the difficulty of identifying successful student learning using these procedures. More and more, school leaders considered how supervision might be a balance between the traditional practice for evaluation purposes and a process that promoted excellence in teaching focused on student outcomes. Attempts to address such concerns are contained in the four goals of the NEPF.

Expectations for using the system of the NEPF are based on a continuous learning cycle (Figure 5). Elements include self-assessment, goal setting, pre-conference, direct observation, post-conference, evidence collection, mid-year review, student achievement, and the final summative rating and narrative (NDE, 2015). The number of observations is differentiated, depending on whether a teacher is probationary in the first, second, or third year of service; or post-probationary. Announced observations last a minimum of 20 minutes. Multiple informal observations, announced and unannounced, are recommended (NDE, 2015).
Figure 5. Nevada Educator Performance Framework Continuous Learning Cycle for teachers and site administrators. The cycle begins with self-assessment. Based on self-assessment, educators develop goals and plans to achieve them. During implementation the educator and supervisor engage in a mid-cycle review culminating in a summative rating and written assessment. Adapted from Nevada Department of Education, 2015.

The NEPF system consists of two categories: Educational Practice and Student Performance. For the 2015-2016 school year, there was no Student Performance included; therefore, one hundred percent of the evaluation was based on Educational Practice. According to the Nevada State Board of Education minutes (July, 2015), Student Performance would be included at 20% in 2016-2017 (10% statewide assessment data and 10% district measures), growing to 40% in 2017-2018 (20% statewide assessment data and 20% district measures). The category of Educational Practice consists of two domains: Instructional Practice and Professional Responsibilities. Each domain contains five standards. Each standard is broken down into three or four
indicators. Every teacher will be evaluated on every standard and indicator every year (NDE, 2015). Figure 6 displays the Teacher Standards of the NEPF. Listed are the Standards for the Instructional Practice and Professional Responsibility domains.

<table>
<thead>
<tr>
<th>Nevada Educator Performance Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Standards</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Educational Practice</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Instructional Practice</strong></td>
</tr>
<tr>
<td>1. New Learning is connected to prior learning and experience</td>
</tr>
<tr>
<td>2. Learning tasks have high cognitive demand for diverse learners</td>
</tr>
<tr>
<td>3. Students engage in meaning-making through discourse and other strategies</td>
</tr>
<tr>
<td>4. Students engage in metacognitive activity to increase understanding of and responsibility for their own learning</td>
</tr>
<tr>
<td>5. Assessment is integrated into instruction</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Professional Responsibilities</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Standards</strong></td>
</tr>
<tr>
<td>1. Commitment to the school community</td>
</tr>
<tr>
<td>2. Reflection on professional growth and practice</td>
</tr>
<tr>
<td>3. Professional obligations</td>
</tr>
<tr>
<td>4. Family engagement</td>
</tr>
<tr>
<td>5. Student perception</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Student Performance</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>(Not included for the 2015-2016 School Year)</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Figure 6. Nevada Educator Performance Framework Teacher Standards. The system consists of two categories: Educational Practice and Student Performance. For the 2015-2016 school year, there was no Student Performance included. The category of Educational Practice consists of two domains, Instructional Practice and Professional Responsibilities. Each domain contains five standards. Each standard is broken down into three or four indicators. Adapted from NDE, 2015.
Now that teacher evaluation, with student achievement included in the rating, has passed into law in Nevada, all districts are required to use the system. Teacher ratings will be reported to the state to develop a statewide picture of strengths and areas of need.

In part, for school improvement efforts to be successful and sustained, the mental models of school leaders must be transparent and aligned to the initiative (Chrispeels et al., 2008; Hannay & Earl, 2012; Keskinikiliç Kara & Ertürk, 2015; Ruff & Shoho, 2005). This study examined the mental models of sixteen school principals to cast light on their beliefs, attitudes, assumptions, and knowledge as related to teacher evaluations. To gain insight into the world of teacher evaluation from the principal’s perspective, the qualitative method of Constructivist Grounded Theory (CGT) was used. According to Charmaz (2014),

Data form the foundation of our theory and our analysis of these data generates the concepts we construct. Grounded theorists collect data to develop theoretical analyses from the beginning of a project. We try to learn what occurs in the research settings we join and what our research participants’ lives are like. We study how they explain their statements and actions, and ask what analytic sense we can make of them. (p. 3)

Constructivist Grounded Theory provides a framework and a process for how to discover and make sense of the hidden world of the principals’ thoughts and actions regarding teacher evaluation.
Chapter Three: Methodology

The purpose of this study was to explore the mental models that principals held regarding the reasons for and use of a new teacher evaluation system in four school districts in the State of Nevada. This chapter presents the methodology used in the study and begins with identification of the research questions and presentation of the research design. Next, the setting and context of the study are outlined, describing the characteristics of the districts, schools, and participants. Following, the reader will find an explanation of the data collection and analysis processes, concluding with the role of the researcher.

Research Design

The research questions guiding this project were:

1) What mental models (belief systems, assumptions, attitudes, and knowledge) do principals hold about the role of teacher evaluation as a vehicle for change in schools; and,

2) Do different school contexts (elementary, middle, and high school) influence the mental models that site administrators hold relative to teacher evaluation?

Qualitative researcher Kathy Charmaz (2014) stated, “We do not exist in a social vacuum” (p. 234). Instead, each person views the world through the lens developed during a life of experiences and training. This study examined the mental models of school principals who were responsible for implementing a new process of evaluating teachers. The study considered the belief systems, assumptions, attitudes, and knowledge of selected principals who had context for describing the process of conducting and
writing teacher evaluations. Therefore, a constructivist grounded theory approach was the basis of the research design of this study.

Constructivist grounded theory considers that both participant and researcher have a role in constructing meaning from the data. This approach is in contrast to objectivist grounded theory which assumes that the researcher is a neutral observer and external authority and that theories emerge solely from the data themselves (not the interpretation of the data). According to Charmaz (2014), “A constructivist approach theorizes the interpretive work that research participants do, but also acknowledges that the resulting theory is an interpretation. The theory depends on the researcher’s view: it does not and cannot stand outside of it” (p. 239). Thus, the researcher brings her experiences and knowledge to the data collection and analysis experience, reflexively and iteratively reviewing interview information during the data collection journey, and adjusting questions and theories as they emerge throughout the process.

Figure 7 below provides a visual representation of the recursive cycle of the constant comparative method of data analysis used in a grounded theory. The constant comparative process informs theoretical sampling and saturation of categories, finally moving to the write-up of the data analysis. Simultaneous memo writing occurs throughout the process (Charmaz, 2014; Glaser, 2013). This model was adapted from “Constructing Grounded Theory,” by K. Charmaz (2014, p. 18). The data collection and analysis procedures in this study were guided by this model.
**Setting and Context**

The study took place in four non-metropolitan school districts in the State of Nevada in the fall of 2016. Though distances were great in that area, relationships tended to be close. Communities were cohesive and the educators tended to come from the area and stay for a long period of time. The school districts together encompassed 7,786 square miles (United States Census, 2010) and served 25,407 students Kindergarten through twelfth grade (Nevada Department of Education, 2015). The student population
range varied from 3,488 to 8,065 students served in each district (Nevada Department of Education, 2015).

Through education legislation, Nevada is divided into three regions, each of which is served by a state-supported regional professional development program. The districts included in this study were in the same professional development region and the principals had received similar in-service training related to teacher evaluation. The total number of schools in the four districts was fifty-two: 25 elementary schools, 13 middle and intermediate schools, 12 high schools, and 2 youth court schools. The two youth court schools were not part of this study because they were not designed to teach the general population of students. The combined student population consisted of 61% White, non-Hispanic; 28% Hispanic; and 11% combined other ethnic groups (Nevada Department of Education, 2015).

**Schools.** The study included sixteen schools. Four high schools were represented, one from each district. The student population range of the high schools was from approximately 960 to 2,250, grade levels nine through twelve. Four middle or intermediate schools were represented, one from each district, with a student population range of approximately 360 to 1,050, grades six through eight. Eight elementary schools were represented, two from each district, with student populations ranging from approximately 360 to 612, grade levels Pre-Kindergarten through six.

Schools included in the study were chosen from 45 schools that received a two, three, or four star rating during the last school rating period of 2014-2015, as determined by the Nevada School Performance Framework (NDE, NSPF, 2015). The NSPF was approved in July 2012 by the United States Department of Education (USDOE) as the
new school classification system in Nevada. The NSPF replaced the previous NCLB system of Adequate Yearly Progress (NDE, August, 2015). Star ratings were paused for the 2014-2015 and 2015-2016 school years due to a change in state assessments. Schools that received a one or a five star rating were not included in the study in order to capture the most realistic pictures from the most prevalent school situations.

The sixteen schools represented a range of socio-economic levels based on Title I funding. Schools received federal funding according to the percentage of children from low-income families attending the school. Six schools were Title I served (receiving funding), five schools were Title I eligible (high levels of poverty but not receiving funding), and five schools were non-Title I. Table 1 represents the distribution of schools receiving Title I funding.

Table 1

Designation of Title I Schools

<table>
<thead>
<tr>
<th>School Type</th>
<th>Elementary</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title I Served</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Title I Eligible</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Non Title I</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Participants. Sixteen principals who were responsible for teacher evaluation at their school were recruited from the forty-five possible participants. The goal to enroll
principals of two elementary schools, one middle or intermediate school, and one high school per district was met. Participants were recruited on a first come, first served basis to fill each available slot. Assistant principals were not included. Only administrators who were serving one school were included.

Of the sixteen principals, seven were female, nine were male, and all were White. The number of years each principal had been in education spanned from fifteen to thirty-four years. The number of years each principal had conducted teacher evaluations ranged from four to twenty-one. Each principal typically completed between thirteen and thirty-four certified evaluations each year (teachers, deans, or assistant principals).

The names of the school principals were identified from publicly available district and state sources. Superintendents of the four school districts were contacted to apprise them of the study. As each school district had its own approach to approving studies, district policy was followed accordingly during the enrollment of the principals. Upon district approval, an email was sent to principals of schools who met the criteria, informing them that they would be receiving a personal phone call with an invitation to participate in the study. To ensure consistency, a script was followed for the invitational phone call (See Appendix A).

Through this phone invitation, the sixteen interviewee slots were filled as each principal agreed to participate. The attempt to maintain a variety of school-level principals (elementary, middle, and high) was successful. Once the administrator agreed to the interview, an email was sent confirming time and place, topic of the study, and information that the interview would be recorded.
Data Source and Collection Procedures

This study was conducted under the auspices of the University of Nevada, Reno, Institutional Review Board (IRB). The method of data collection in this study was the interview. An intensive interviewing process was used with a semi-structured interview protocol as guide. Intensive interviewing “typically means a gently guided, one-sided conversation that explores a person’s substantial experience with the research topic” (Charmaz, 2014, p. 56). The aim of the intensive interview was to allow the participant flexibility to answer questions and to attend to subjects that were important to his or her context for understanding. The intensive interview also allowed the researcher to gently probe and delve deeper into certain topics that arose during the interview, thus pursuing key theoretical concerns (Charmaz, 2014).

The flexibility of the intensive interview supported the development of knowledge related to the mental models that principals held by co-constructing an understanding of each principal’s experiential reality. The structure of the interview was based on pre-developed questions with a broad range of possible follow-up questions as the conversation dictated (See Appendix B). This approach offered benefits to both researcher and participant in terms of flexibility, honoring participant input, and collecting data. This type of data had a higher potential to accurately portray the mental model of the participant as opposed to a structured interview or a survey. An exploration of the personal experiences of the participant contributed to understanding why respondents held their mental models. The flexibility to guide the conversation in such a manner allowed the insights of the participant to emerge (Charmaz, 2014). The researcher spent time prior to the actual interviews practicing with administrators not included in the
study to refine interview questions and interviewing techniques. Figure 8 outlines the benefits for both interviewer and interviewee of following the constructivist grounded theory intensive interview approach.

<table>
<thead>
<tr>
<th>Intensive interviews allow an interviewer to…</th>
<th>Intensive interviews allow research participants to…</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Ask for an in-depth description of the studied experience(s)</td>
<td>• Break silences and express their views</td>
</tr>
<tr>
<td>• Stop to explore a statement or topic</td>
<td>• Tell their stories and to give them a coherent frame</td>
</tr>
<tr>
<td>• Request more detail or explanation</td>
<td>• Reflect on earlier events</td>
</tr>
<tr>
<td>• Ask about the participant’s thoughts, feelings, and actions</td>
<td>• Be experts</td>
</tr>
<tr>
<td>• Keep the participant on the subject</td>
<td>• Choose what to tell and how to tell it</td>
</tr>
<tr>
<td>• Come back to an earlier point</td>
<td>• Share significant experiences and teach the interviewer how to interpret them</td>
</tr>
<tr>
<td>• Restate the participant’s point to check for accuracy</td>
<td>• Express thoughts and feelings disallowed in other relationships and settings</td>
</tr>
<tr>
<td>• Slow or quicken the pace</td>
<td>• Gain a new perspective on past and present events</td>
</tr>
<tr>
<td>• Shift the immediate topic</td>
<td>• Receive affirmation and understanding</td>
</tr>
<tr>
<td>• Validate the participant’s humanity, perspective, or action</td>
<td></td>
</tr>
<tr>
<td>• Use observational and social skills to further the discussion</td>
<td></td>
</tr>
<tr>
<td>• Respect the participant and express appreciation for his or her participation</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 8. Benefits of the intensive interview approach for both the interviewer and the interviewee. Both parties participate in co-constructing a realistic picture of the explored topic. (Adapted from Charmaz, 2014, pp. 69-70).*

The open-ended questions prepared in advance provided an entry point for the conversation. However, in contrast to the objectivist grounded theory approach, where the interviewer asks the same questions of all participants before analyzing the data, the intensive interview allowed for choosing follow-up questions during the interview and
obtaining more detailed responses as the moment presented itself. Interviews were analyzed after each episode, providing guidance for the next interview during an iterative process of analysis, rather than gathering all interviews for a collective analysis at the conclusion of all interviews.

The interview process. Interviews were conducted on the school site in the principal’s office or other private conference room. Each interview lasted between sixty and ninety minutes. The interview process followed a general protocol: Introduction and demographics, initial open-ended questions, intermediate questions, and ending questions (Charmaz, 2014). To establish rapport and provide context for the conversation, an introduction including the goals of the study was shared with the participant along with reasons for audio-recording the interview. Interviews were audio-recorded to maintain accuracy and to allow the researcher to listen carefully without the necessity to develop extensive notes during the interview. A consent form was offered at this time (see Appendix C). During the introduction phase, questions covered demographics and built background such as: What has been your journey to get to this place and time? How long have you been in your current position? What previous positions and locations (teaching, administration, etc.) have you occupied?

Next, the interviewer moved toward initial open-ended questions to ease the principal into recounting familiar information and developing context. This conversation was designed to provide information leading to understanding the beliefs, attitudes, and assumptions of the participant. Questions were asked such as: How long the have you been responsible for doing teacher evaluations? What have been your main influences? Who have been your main influences?
Following the initial open-ended questions, the intermediate phase moved to the current teacher evaluation process to ascertain information about assumptions and knowledge as well as beliefs and attitudes. Questions related to the administrator’s training in evaluation, the new evaluation roll-out in his or her district, and his or her understanding of the process and rubrics. Other areas explored included strategies that helped the participant manage the requirements of the system, thoughts about the most important aspects of teacher evaluation, the purposes, and what the principal did with the information that he or she acquired during the process. Changes made in the principals’ thinking, beliefs, or processes were also explored during this phase.

The final phase of the interview consisted mainly of questions to conclude the conversation on a positive note. However, it was not unusual that this was the phase that rendered the most interesting stories and greatest revelations around beliefs, attitudes, and assumptions. Questions were used such as: How have you grown as a school leader? What advice might you give a new evaluator regarding this process? And, is there something that you might not have thought about before that occurred to you during this conversation? At the conclusion of the interview, time was reserved to allow the principal to add to or clarify any comments or to ask questions of the researcher.

**Data Analysis Process.** Analysis began with the first interview transcription. The first interview was typed by the researcher and general thoughts were recorded through memoing to capture possible theme trails early on in the interview process. The transcript was then read carefully and each line was coded using the gerund form of the verb. For example, if an interviewee stated, “I am overwhelmed by this process,” the line was coded “Feeling overwhelmed.” This process allowed thoughts, perceptions, attitudes,
beliefs, and knowledge to be captured in a way that revealed themes suggested by the respondent rather than fitting the data into pre-determined themes. These authentic themes brought focus to what was happening in the data. Once the interview transcript was coded, the codes themselves were read and preliminary groups of ideas were identified. This process was repeated for the second interview and the codes were compared with the first interview to further identify groups of common ideas. The third interview was coded and compared in the same way. After three interviews were completed, some general themes were beginning to emerge.

The general themes from the analysis influenced to some degree the questions used in subsequent interviews. Though the same general interview format was followed, more time and emphasis were given to questions that allowed the principals to expand on their thinking. Questions were revised to elicit clarification of possible themes. One such question was, “If you were to explain your philosophy around teacher evaluation, what might that sound like?”

At this point in the process, codes were recorded on sticky notes and hung on large posters, grouped by the general themes. The fourth and subsequent interview transcripts were then processed in the same way as the first three until all of the transcripts had been typed, read, coded, and compared with previous interviews. Along the way, codes were grouped and regrouped as themes were uncovered, revised, or phased out. Memo notes were also considered to confirm or challenge the themes. Once it became evident that no new themes were emerging, the themes were categorized under major themes.
The next step entailed conferring with two advisors from the doctoral committee. Input and questioning from the advisors ensured that bias from the researcher was avoided and fresh eyes encouraged fresh ideas in analyzing the data. The themes were regrouped several times during this phase, arriving at the first general plausible mental model. As the data were recategorized, it was determined that the first model was not sufficiently explanatory of all the data. The mental model was revised, again with input from the doctoral committee advisors, and the final representative major themes were identified.

Reorganizing and considering the themes from different perspectives resulted in a mental model that represented the natural flow of the respondents’ experiences and answered the research questions. During the process of writing the results, themes were revised and polished through discussion with advisors. Chapter Four describes the results.

**Role of the Researcher**

An ethical consideration in regards to this study was the relationship of the researcher to the participants. A number of years ago, the researcher was employed as a teacher and staff developer in one of the districts represented in the study and was acquainted with some of the administrators from that time period. In addition, the researcher was employed at the Department of Education and subsequently served as a site administrator and as the director of a professional learning resource in the region where the four districts were located. Through these roles, the researcher had possible professional encounters with the administrators in the study through support services and the delivery of professional development. Every effort was employed to follow a consistent interview protocol so that each participant had the opportunity to respond on
an equal footing. The practice of bracketing was employed during data collection and analysis to consciously recognize and set aside assumptions and preconceptions.

Guidance was procured from doctoral committee advisors during the analysis process to hinder bias in the interpretation of results.
Chapter Four: Results

Results from the analysis of 16 principal interviews are presented in the following chapter. Interview questions revolved around principals’ beliefs, attitudes, assumptions, and knowledge regarding a new teacher evaluation system. Analysis of the data revealed three major themes. The first major theme was Tension. The second major theme was Desperately Trying to Figure It Out. The third major theme emerged as Glimpses of Resolution. These three themes, along with associated subthemes, are presented.

Tension

All principals described the Tension that they were feeling as a result of the dilemmas they faced on a daily basis. They wanted to maintain the integrity of the purpose of evaluation, but it seemed they could not carve out the time to do the process justice. One principal explained it quite simply, “I think evaluations are so important and I just wish we had the time to always do them right.” The main source of this Tension was described by the majority of the principals in terms of two subthemes: 1) Importance of Doing Teacher Evaluations, and 2) Feeling Overloaded and Overwhelmed.

Importance of doing teacher evaluation. According to the principals in this study, the purpose of teacher evaluation is to increase teacher capacity to increase student learning. As one principal described it:

The ultimate purpose … is to improve instruction. That is the ultimate goal.

Accountability pieces to make sure – you’ve got to cross your ‘T’s and dot your ‘I’s and all that stuff, yes, we get that, but the ultimate goal should be to improve instructional practices for kids. That's the bottom line, I think, is the students. Improve the instruction.
Principals identified a variety of ways in which teacher evaluation could be used to support the increase of teacher efficacy. All of the 16 principals identified providing evidence-based feedback to teachers as one of the most important roles they fulfill as school leaders. For example, “I've always thought that teacher evaluation was really important because I think people deserve feedback, and they need feedback, if they really want to improve.” Administrators noted that evaluation should not be a “gotcha,” but a way to build relationships and develop a positive culture in the school. As one principal put it:

The more you walk through, the more you sit in their class, the more you know what they’re doing… and teachers are more comfortable and they see you not as ’I’m coming in to get you’ and that’s the culture you want to have, that you are coming in to support them and help them. And I think that’s the biggest thing that through the evaluation process you need to do.

Other ways principals described increasing the capacity of teachers included helping teachers with their self-reflections and learning. One principal noted, “I really feel the best thing is the reflection. I think that an evaluation is only as good as a teacher, what they do, use it for, or what they take away from it.” Self-reflection was identified as a tool to help teachers understand that they were not going to be perfect, and they did not have to be. The goal was to maintain an attitude open to new ideas and input to keep growing and learning the craft of teaching. One principal explained:

It is trying to change their mindset about, this is a tool for us to use to work on portions of your instruction that could be improved because we’re not perfect. And get them out of that mindset that they have to have a 100% or otherwise
they’re not a good teacher. It’s interesting to listen to that, but it’s very difficult to get people to change their mind and to understand that. If you get a lower mark then maybe that’s the thing we’ll look at next year for improvement and work on it. It’s hard to let that go for some people.

Principals reported that evaluation was connected to the growth of the whole school. One principal noted, “Your school improvement plan is the teacher improvement plan.” In other words, for the school to improve, teachers should be improving. Four principals, from two different districts, explained that evaluation information was being compiled by their districts by each evaluation standard and associated indicators and given to the schools. They were using the data as a whole-school profile to inform decisions about professional development, bringing all staff members together to work on a common goal. One principal described:

In addition to that, our district, which is a very cool thing…they mapped all of their teachers and what their scores were for all of the indicators so that they could see if they had trends that they needed to help teachers in terms of professional development. So our district has done that for all of our schools.

All the principals described the process of evaluation as crucial to observing and judging curriculum in action. They indicated that teachers should be providing instruction that is relevant and based on students’ needs. Two principals added that observing for curriculum and instruction was a way to connect with teachers. Two other principals commented that if teachers were using strong, effective instructional practices, then students should be performing well. Principals used the evaluation process to observe whether best practice strategies were in place and to provide support training if they were
not. Said one principal, “it gives you a kind of a roadmap of what we’re looking for in good instruction and a scale to measure it by.” According to the principals, the administrator's role should be to provide support to teachers: providing teachers what they need in terms of feedback; access to implementation specialists or learning strategists; materials, books, or other information; or the opportunity for in-service training. A number of principals indicated that if teachers were open to considering suggestions and growing, the “final evaluation piece should be good.”

Virtually all of the principals noted that the evaluation system should be a coaching system where there are deep discussions between teacher and administrator. One principal described it in terms of a philosophy:

My overall philosophy is that the evaluation system is a coaching system and the evaluative piece at the end, if you're doing it correctly and you're doing it consistently…you're having some good discussions and you're making improvements and you're giving them support, whatever they need, their evaluations at the end should be pretty darn decent, with some - maybe a hole here and there… Evaluation system is all about coaching and improving.

A few principals identified evaluation as an important piece to maintaining a positive culture at the school. They explained several ways that this was accomplished. For example, administrators who are consistently in classrooms, talking with teachers, and providing support and feedback, have a better understanding of the pulse of the school; they develop better relationships with the staff and communicate clearly and more often about school issues; they are able to see the big picture and help tie
initiatives together; and, they can connect teachers to others who are working on similar strategies or set up opportunities for peer observation. One principal noted:

It’s hard to evaluate somebody if you don’t know them. I think it’s like kids coming into a classroom. You’ve got to build relationships first, there has to be a trust factor…If there’s not trust, this tool and this evaluation process becomes more difficult because they don’t feel that you’re giving them viable information, it’s more of a ’got you, you’re trying to get rid of me.’

Most administrators affirmed that they were not trying to use evaluation negatively. They described how the process provided a plan of assistance when needed that supports a teacher with specifics. These specifics may help a teacher see and understand the reality of his or her teaching. A few principals pointed out specifically how the rubrics could be valuable to a teacher as a means of self-reflection about his or her performance. The rubric could function as the basis for frank conversations about effectiveness and, if necessary, help a person decide whether to stay in teaching or not. One principal described a situation with a teacher where the teacher used the evaluation rubrics to score herself and eventually voluntarily left teaching. “But I took her, what she said, and then figured it out mathematically with her there, and it was ineffective and she knew it. When we got to the end, she knew. I had let her score herself.”

**Feeling overloaded and overwhelmed.** All 16 administrators agreed on the principles and importance of evaluation. At the same time, they acknowledged that there are important logistics that must be considered. One principal summed up both aspects in the following quote:
Well, I think it's probably two-fold. Again, I’d go and put on my supervision hat first and then say, to me, the purpose of any feedback or observations, assessments that we do for teachers is to give them feedback about what they're doing in the classroom… That's part one and then part two is just to make sure that we meet the requirements of the law because that's really part of what the evaluation process is, is making sure that we’ve followed all the timelines and all the paperwork and all that. Then the first one is way more important than the second one, but you’ll get in trouble for the second one, so you've got to watch out for that.

All principals named the new state evaluation system, the Nevada Educator Performance Framework (NEPF), as a high-stakes and time-consuming additional initiative. The first major concern that all principals expressed was completing evaluations based on all ten standards for all teachers every year. Principals reported personally evaluating between 13 and 34 certified staff (teachers and assistant principals). These numbers did not include classified staff (teacher aides and clerical positions), who must receive evaluations based on a different system.

Interviewees explained that the new system required both administrators and teachers to learn new processes, standards, indicators, and their meanings and nuances. One principal noted his own needs in terms of knowledge:

The other part is, and I’m not a hundred percent there yet, is you’ve really got to know what it is you’re evaluating – you know, what those terms mean, what is the tool you’re using and what does quality instruction look like? It can’t be your
opinion, you know, ‘Because I taught this way before, this is how I teach so that’s the best way.’

Principals explained that the requirement of a minimum of two pieces of evidence for each indicator (thirty-four indicators in all; nineteen in the five standards of the Teacher High-leverage Instructional Standards and fifteen in the five Professional Responsibilities) caused anxiety among most principals and teachers. They reported struggling with a number of unresolved questions such as how to collect enough evidence for all indicators, defining what constituted evidence, and being able to “recognize it in action.” Most principals recounted concerns with “fitting it all in.” One principal pointed out size and time concerns with the new document as he was “learning while doing”:

The NEPF has been a steep learning curve because it's very large; the document itself is very large, with all the indicators. And last year was - I don't know how much time I spent on observation, coming back and looking at what I scripted and going, Where does it fit? because I didn't know the document well enough because we were getting trained on it. We were going through it…So learning the document was…a steep learning curve.

A handful of principals noted that teacher evaluation had always been part of the daily routine of the school administrator, but that because the NEPF was new extra time was required to learn the specifics and nuances of the rubrics. These principals noted that, though the new evaluation system was overwhelming, there were other causes for being overwhelmed as well. One principal stated, "It's not [evaluation] that's overwhelming me. It's more the new people... coming in ill-prepared and then all the sudden Read by Three
literacy stuff... A lot of things hit at the same time. It's not NEPF that's creating the workload, it's the other things." Another principal explained:

We’re not working harder because of NEPF, we're doing almost the same things.

We're working harder because of other initiatives. It's pretty hard to even get that through to my colleagues. NEPF could make our jobs easier if we do it right. It's the other things that are giving me the headaches at night.

Without exception, principals proclaimed that the volume of new tasks and timelines added an increased sense of pressure to their evaluation responsibilities. Meetings for teacher evaluation had become more numerous. Principals described how the new system required the same expectations for all teachers every year: a complete evaluation process on all ten standards. These expectations included setting instructional and student learning goals with teachers, increased numbers of observations with completion timelines, pre- and post-conferences for formal observations, and an added mid-year review. Principals all described the final write-up of the evaluation in the spring to be very time consuming. Time needed was reported as anywhere from two to twelve hours per evaluation. “I still think in some ways I'm dreading the final evaluation again,” intoned one principal, “just because it is so time consuming and making sure we get those done appropriately. That is my biggest fear right now, still trying to go back in [to classrooms to collect evidence].” Several principals reported using weekends, after school, and/or all of spring break to complete the write-ups. One principal mentioned that he and colleagues had had to take time away from campus to complete the write-ups as there were too many interruptions during the school day. However, spending time away from campus created other issues, so it was described as an uncomfortable trade-off.
Principals expressed the need for clearer communication from the state level in regards to timelines and due dates. Templates seemed slow in appearing, causing double work in transferring information, particularly toward the end when evaluations were due. One principal was frustrated because earlier there were too few specifics but now there were too many:

I think not having a lot of specifics and a lot of process when I first came into administration for the first couple years, it was not really helpful. Where we are now, on the other end, I think in some cases is too limiting and too all encompassing, there's too much to it.

Additionally, principals were frustrated with teachers not seeing the benefits of the self-reflection or equating the expectations of the rubrics with raising learning opportunities for students. Said one principal, “I think that is the frustration that I have, that not everybody is here for that common purpose of kids and have that vested interest in our kids and that’s what frustrates me more than anything.” Another principal expressed, “I feel like I’m saying the difficult things but it’s not encouraging the desire…I want to encourage the desire to change. I want people to want to change if they need to.”

A few principals expressed anger at the way the final expectations for implementation were rolled out. Though every principal had heard of the system and had received some basic training in the process and the rubrics, most agreed that it was not enough and that the timelines were imposed too quickly for effective use. Principals felt there was not enough time to learn the rubrics and to practice using them in observation. Concerns were expressed about data collection, reliability, and understanding the new
scoring system and high-stakes expectations. One principal described colleague reactions as follows.

   It was sick around the whole state, it really was. We went up to NASA, up there in June. Principals were mad, teachers were mad and it’s kind of like, “I think you guys are all overreacting to this.” I had probably listened to three principals say, “This is going to drive me out. I’m quitting because of NEPF.”

Another principal also expressed the viewpoint of colleagues:

   I think because some of my colleagues had been in the district a lot longer than I have and they’re used to Danielson so, “that’s the way it’s been, that’s what I’m comfortable with and that’s what I like because I can schedule them out, it’s done.” Switching to this is just such a lengthier process that, I think, that’s part of the negativity that could be felt.

Many principals described the pressure they felt from what they considered to be imposed, unreasonable timelines. They expressed being in constant fear of missing deadlines. For example, initial observations must be completed for probationary teachers within the first forty days of instruction. These observations are required for teachers in their first, second, or third year of probation. One principal reported twenty-five probationary teachers at the school and expressed consternation at the inflexible timeline. At the same time, pressure to “do a good job” on the final evaluation was high to create legally defensible evaluations based on the new scoring system.

   It was not only the length of the written evaluation document, but the language that caused frustration. All principals described difficulty becoming proficient with the new rubrics. For example, they craved more clarification and concrete examples of the
indicators. Many principals described confusion among principals and teachers regarding the meaning of the indicators and new vocabulary. This lack of understanding was seen as a barrier to being able to score an observation accurately and consistently. Because of the high-stakes nature of the final evaluation, and the necessity for increased specificity in evidence collection, scoring, and feedback, principals indicated that a deeper understanding of the rubrics was crucial. Phrases like metacognition, most versus all students, high-cognitive tasks, and discourse seemed vague, unclear, inconsistent, and/or had different meanings for different principals. Without a common interpretation of what such concepts “looked like in the classroom,” evaluations could be questioned, contested, or legally challenged. One principal opined, “The terminology is like learning a foreign language. All the terminology is different. Sifting through that and tying it to what you see in the classroom and the practices, that's what's taken some time, and trying to build in some inter-rater reliability.”

Principals described routines and knowledge they already had in place regarding teacher evaluation and many had spent a great deal of time honing their skills and understanding of previous systems. Three models had been used: the Danielson model, based on the work of Dr. Charlotte Danielson (2007); the Marzano Focused Teacher Evaluation System (Marzano, 2011); and T4S (Teach 4 Success) (WestEd, 2009). All three approaches had involved training and deeper understanding of what to look for in the classroom. Principals struggled to find commonalities and similarities with the NEPF when it was introduced. According to most principals, though there were some crossover concepts, the new system required a different level of understanding and knowledge around instruction and learning that had not been previously emphasized. One principal
pointed out, “Just mentally, through my processing, I’m not looking for the things that would have been commonplace before for you to look for. Those have kind of slid off to the sides for right now.”

Principals expressed feelings of loss regarding the former processes and tools that represented so much of their knowledge base and the time spent to learn them. Many felt their previous knowledge was now obsolete and were struggling to regain a sense of control and efficacy over the new material. They were caught in the transition between what they used to know and do and new expectations.

In a few cases, principals expressed a feeling of loss of control tinged with fear and anxiety, which added to the feeling of being overwhelmed. “Are they going to change this again from the state level? I will just get this learned and they will change it,” were among principals’ remarks. Another fear that surfaced was one of “doing it wrong.” Because the process was so new and had had little practical vetting, principals were concerned that they might complete the process incorrectly, causing repercussions from the staff, district, or state.

Another overwhelming new requirement was looming for the 2016-17 school year: the upcoming addition of student data to the evaluation. In 2016-17, student achievement data was included in the NEPF at a total of 20% of the overall evaluation: 10% based on State testing, 10% stemming from Student Learning Goals. The Five High-leverage Instructional Standards counted for 60% and the Five Professional Responsibilities Standards counted for 20%.

State achievement testing data were to be calculated for each school and translated to a score of 1 – 4. Every school would then receive a 1 – 4 score that was
included at 10% of all certified teachers’ and administrators’ evaluations. Additionally, each teacher and administrator would receive another data score based on the success of his or her Student Learning Goal (SLG), which was scored on a 1 - 4 rubric also. The 1 – 4 score was then included at 10% of the final evaluation. The 2016-17 school year was the first time teachers or administrators in Nevada would have these types of data tied to their evaluation. Because of the newness and uncertainty around this process at the time of the interviews, a number of principals expressed anxiety about the additional scoring of the SLG. Scoring evaluation indicators and SLGs were new tasks that principals must perform in completing the final evaluation. Many principals expressed concern about being both coach and judge for the new requirements.

The evaluation process was not the only reason principals expressed feeling overloaded and overwhelmed. They noted several new areas requiring attention: New curriculum and materials were being adopted which required professional development and extra time to support staff, especially in the areas of mathematics, science, or writing; new areas of focus were required by the State in terms of grants approved through legislation. Every principal named at least one initiative that they were juggling. The list included: STEM (Science, Technology, Engineering, and Math), Read by Grade Three, Victory Schools, Zoom Schools, Bullying requirements, Infinite Campus, Smarter Balanced, and End of Course testing. All of the principals interviewed lamented the number of initiatives in progress at their schools and the time required for new learning and monitoring.

A number of principals identified teacher shortages as an additional distraction because of many long-term substitutes or positions filled from the Alternative Route to
Licensure (ARL) system. Substitutes and ARLs required extra attention due to their lack of training and experience.

**Desperately Trying to Figure It Out**

The identified Tension between the Importance of Doing Teacher Evaluation on the one hand and Feeling Overloaded and Overwhelmed on the other seemed to be a catalyst for principals to choose their next steps. Principals described balancing their time, existing skills and knowledge, and beliefs with the new evaluation system. They could not just focus on one, all had to be addressed. Therefore, principals were Desperately Trying to Figure It Out. As administrators recounted their efforts, two subthemes were found: It’s a Catch-22 System and Getting a Handle on Time Management.

**It’s a Catch-22 System.** Each administrator in this study had a different set of skills and ideas to manage time, collect data, conduct meetings, and complete paperwork. Managing the system was reported as especially difficult by elementary principals who did not have assistant principals or deans to share the load of meetings, observations, and data input. Setting up goal meetings, pre- and post-conferences in conjunction with formal observations, writing up feedback, inputting evidence into a summary document, mid-year review meetings, and eventually the final evaluation write-up, became so cumbersome that principals felt the system worked against itself. For example, one principal reflected:

Like in March or something, I spent large chunks of time doing it. It got in my way of doing it. Okay, I met with them and I tried to input stuff, then I met with them again and input more stuff, then it felt like I was just always doing the end-
product, as opposed to just observing and taking notes on observations and gathering it.

Another principal speculated that evaluating all ten standards for every teacher every year would simply lead to repetition. She wondered, “I think it's an important part, but I don't know if it's realistic. At some point, is it just going to be the same thing over and over?”

Many principals pointed out that doing the same thing over and over would devalue the system and it would lose its power to focus on instruction and encourage change. Another principal was quite clear about the underlying reasons for introducing the new system and described how it was defeating its own purpose:

But, you know and I know, the whole impetus for this whole new evaluation was we needed to have this big comprehensive evaluation tool because we had too many mediocre teachers out there. But they [Department of Education/legislators] just shot themselves in their own foot. Because they made it too unmanageable for us to use it the way it needs to be used to actually work with those teachers that need the most work…we all know who our weak teachers are…so give me an opportunity to focus on that teacher and not spend 15 hours for my other teachers that are flying and doing a great job so that I can check all the boxes and say that I’ve been compliant with your tool. This teacher needs the time. This teacher needs me to go into his or her classroom 20 times instead of the two that I’m able to go do because I have to be in everybody else’s classroom that many times.

Principals seemed to feel that they had to circumvent the system in order to “do” the system: it was a Catch-22. “Doing the system” the way it was intended detracted from
areas that principals felt they needed to focus to be able to make the system work to improve instruction to improve student learning.

Principals were concerned that the completed evaluations be well-written, based on clear evidence for scoring, and provide useful written feedback. Ironically, most principals reported that teachers were not inclined to read the final evaluation. Though a couple of principals expressed that teachers seemed appreciative when reviewing the final evaluation, the main attitudes principals encountered from teachers were, “I feel sorry for you that you had to spend so much time writing this,” or, “Did I pass? Did I do ok? Great, where do I sign,” or “Just tell me the bottom line – What is my number and what do I have to do?” One principal summed it up in this way:

Actually, I don't know that they reflected on it as much as the administrator who had to write it. So, I hate to admit that [laughs] because I think I did. I think I spent a lot more time on it than they did.

**Getting a Handle on Time Management.** Principals described a deep sense of responsibility for their schools, their teachers, and their students. They recounted a wide variety of approaches to putting structures and processes in place to support their charges. Principals wanted to make sure they were both “doing this thing right and doing the right thing.” Conceptually and operationally, principals were at different stages in the process of trying to figure out this system. However, they appeared to be trying to maintain the integrity of evaluations and, at the same time, fulfill the new requirements.

Scheduling of required activities was a concern among all principals because, even with the best intentions, school issues and events interfered with teacher meetings and classroom visits. The urgency of dealing with discipline, parents, and other daily
occurrences made keeping to a schedule difficult, even with the evaluation process a priority. For example, administrators described the requirement to attend Individual Education Plan (IEP) meetings for students with special needs; thus, it was necessary to plan around those events. Elementary principals in particular explained how preparation and professional learning time was frequently scheduled so that teachers could meet collaboratively. Scheduling time with teachers one-on-one for evaluation-related conferences took away from planning and collaboration time. One principal encapsulated many of the issues in the following description:

I think [teacher evaluation] is truly one of the most important things on the principal plate or the administrative plate to make sure we do an honor and spend time with. And it is probably one of the easiest things to have not go that way because it's not necessarily a day-to-day; you might have a week go by. Like last week we didn't do any observing because the kids were here until 12:40, it was parent-teacher conferences, and report cards, and end of the quarter, and all of that, so we didn't schedule anything. It's easy for another day to go by and, well, I didn't get to that.

Proactive priority scheduling was one of the primary approaches reported by principals; interviewees shared a variety of strategies. Some held group goal-setting meetings where teachers collaborated around setting grade level, department, or content goals with the administrators as facilitators and advisors. Some tapped teacher leaders or external coaches to lead goal-setting. Many principals used email or internet resources such as Google to share information regarding self-evaluation, goal-setting, pre-conferences, or post-observation notes and feedback.
A few principals utilized pre-printed hard-copy notes with indicators that could be checked off and left in the classroom with space for added comments. Two principals challenged themselves to leave a certain number of these notes per day as feedback for teachers. This practice was envisioned to ensure that administrators got into classrooms to collect evidence on the standards. Several principals mentioned sending written notes by email after informal observations and allowing strong teachers to decide whether they wanted to meet face to face. A few principals described sending mid-year review notes in advance to help shorten the face to face mid-year review meeting if there were no concerns about teaching. A couple of principals described using video as a way of sharing classroom events. A few principals described collecting information from teachers through portfolio-type binders, but there was disagreement as to whether this strategy was helpful or not. One principal described the results of using binders as follows:

I made them artifact binders and made them bring them to the meeting. I don't know that that I saw anything - they were more like brag books for them. I mean, they wanted to show me their entire binder [laughs]! I think I got more from just observing when I was in the room, whether or not I thought they were meeting what they needed to meet, than with looking at papers afterwards.

Many principals described spending hours going through binders when writing the final evaluation or sitting with teachers as the teachers shared their collection of evidence. Though there were some pearls in the evidence, most administrators considered this an inefficient exercise; they were learning how the evidence could be better managed.
In terms of organization, principals tried different approaches to stay organized and fit all the requirements of the formal evaluation into the school-day schedule. In many cases, principals shortened the required pre-conference time by sending questions in advance or asking for lesson plans by email. This way, the time in the face to face pre-conference was shortened and used to just make brief clarifications. A handful of principals tried to conduct pre-conferences by email. For informal observations, pre-conferences were normally not held at all. Some principals reported also shortening post-conferences by sending written feedback to the teachers in advance. Teachers were able to prepare their thoughts, responses, or questions beforehand. Often, the written feedback was enough and the teacher could indicate that they did not need a follow-up face-to-face meeting.

In terms of their own learning, most principals reported that collaborating with colleagues was the preferred approach to engaging in deeper discussion about the NEPF process and standards. Administrators described learning as a site team with an assistant principal and/or dean. This practice gave some principals the opportunity to observe classrooms with others at their own school and to discuss findings with the site leadership team. Using colleagues to debrief the process and the content provided an opportunity to share how written feedback was given and/or how narrative write-ups might be crafted. One principal described scheduling eighteen weekly meetings over the course of the year with the administrative team. Standing items on the agenda were who they had observed that week, what they had seen, and what kind of feedback was provided. Additionally, the principal described collaborative evaluations to calibrate evaluation writing within the team:
This is how I’m writing this feedback, what do you think? That kind of thing, so we help each other out that way. And then last year, actually the last two years, I’ve done collaborative evaluations with my new folks for one of their higher needs people so that they have that.

Three principals described meeting with other administrators from their area to visit each other’s schools and observe in classrooms. They discussed the teaching and calibrated the scoring together based on the evaluation rubric. They also helped each other understand how the indicators in action might look different at different school levels (elementary, middle, high school).

Four principals recounted participation in district-organized “learning walks” over a series of four meetings, organized in collaboration with the Regional Professional Development Program. In that model, all district administrators met for a brief professional learning conversation. A pre-determined focus was announced and the administrators were divided into teams, different for each meeting. The teams then visited a school where they observed and scored three classrooms together. Upon return to the meeting area, the observations were debriefed in a facilitated professional learning community format. The four principals who participated in this activity reported that it was useful in developing respect among colleagues, in supporting a learning team culture, as well as in deepening learning about the evaluation rubrics.

At least half of the administrators described enlisting teachers as learning partners. Using the evaluation rubric as a talking point to guide the post-conference conversation, teachers were able to give insight into their interpretation of vocabulary and what certain concepts looked like in the classroom. This approach allowed the rubrics to be the
“expert criteria” instead of basing the conversation on the administrator’s opinion. These administrators indicated that creating feedback with the teacher was a learning experience for both teacher and principal as they negotiated what each one saw and experienced in the classroom.

Developing supports among teachers was another approach. A few principals described enlisting professional learning coaches to help teachers deepen their understanding of SLGs or instructional concepts contained within the NEPF. Others described professional learning community conversations where teachers shared instructional strategies with each other that addressed NEPF standards such as metacognition.

Principals described experimenting with an assortment of ways to collect data and found some more efficient than others. Most interviewees reported beginning with what they knew, such as the traditional “scripting of everything,” which proved to be very time consuming. Principals noted that this method worked well for information gathering when the new system was first rolled out. Because they were unsure of what they were looking for in the beginning, with this method they were able to go back to the office to compare notes with the evaluation rubrics. However, most principals indicated that they began opting for more targeted, short observations of ten to fifteen minutes as they became more proficient with the indicators. Administrators explained how they concentrated observations by 1) determining a focus from the pre-conference for formal observations, 2) visiting several classrooms in a row with the same focus from the rubric, or 3) completing one short formal cycle (pre-conference, observation, post-conference).
followed by several informal observations with a follow-up email and/or face to face conference.

Principals often mentioned how the NEPF system needed to be revisited, that every teacher evaluated every year on every standard was too much. Almost all interviewees suggested options for revision. One idea was to reduce the number of standards to be evaluated each year and to cycle through them on a rotating basis, for example, two instructional and three professional responsibilities in one year followed by three instructional and two professional responsibilities the next. Another option was to forego complete evaluations every year on teachers who were effective or highly effective. It was suggested that these teachers could receive an evaluation every other year with observations only on the “off” years. Many principals mentioned hoping the next legislative session would result in changes.

**Glimpses of Resolution**

Though all of the principals expressed challenges, there were a few Glimpses of Resolution as they navigated the new evaluation system. In spite of Feeling Overloaded and Overwhelmed, most principals showed a great deal of pride in how they were tackling the new challenges. They wanted to share ideas they were trying which seemed to move toward their concept of the important goal of teacher evaluation: increasing teacher capacity to increase student learning.

A handful of principals indicated that observations were shifting from a “dog and pony show” (where teachers planned a specific lesson for the formal observation) to observation of elements of instruction in the rubrics on a more consistent basis. These principals described collecting evidence of teaching and learning according to the new
indicators and providing feedback to teachers. “That’s how we grow,” said one. These principals found the new rubrics to be more rigorous than their previous systems. The rubrics were described as useful to frame delivery of more specific and focused feedback. A couple of principals appreciated the set of Five Professional Responsibilities Standards and mentioned wishing the professional responsibilities had been given more weight in the scoring. One principal declared, “I might be one of those weird outliers…but I felt the NEPF has given me more leverage than any evaluation tool I’ve used in regard to impacting teacher performance.”

Principals described the requirement for more face to face conversations as a positive thing. One interviewee remarked:

Those conversations take place - which is the most important thing that we can do. And after that, then I go in and I make notes on the bottom with, 'Hey, here is what we talked about. These are the things that we saw. These are the things that we could be working on,' and make suggestions or recommendations or just whatever needs to be talked about with feedback to the teacher. It's helped. It’s working well.

One principal recounted an experiment that involved using teachers as in-house coaches for other teachers. A prep period was freed up out of the daily schedule for four strong teachers. Each of the four teachers was required to use the time to meet with colleagues to share ideas about staff-identified interests or needs. These “lab” teachers also began to lead whole staff conversations about deeper understanding of the NEPF standards. The principal depicted the situation as follows:
We're looking forward to that and we know that we've already gotten them in front of our staff. All of our staff meetings this year focused on professional development in the morning and not in a typical sense…we’re not going to deliver information to you but just as in professional collaboration. So they're starting to lead those conversations now and we're kind of taking a backseat to that as administration to support them but to get them in front of those teachers supporting things that they're doing that are meaningful to them and seeing how they can support their peers.

A few other principals explained calling on district level professional learning support personnel or regional professional development program facilitators to work with staff, either as a group, in small groups, or individually. Two principals described teachers participating in rounds of peer observations to provide opportunities for strategy sharing. Another principal described setting up mentor teachers for new teachers in the building to help them learn the new evaluation system.

Other examples that fell within the Glimpses of Resolution theme were more individualized to each principal. One principal related how she began publicly honoring (for example, at staff meetings) samples of the new indicators that she had seen in action and how she was allowing more choice in focus goal areas. Student learning goals were identified by two principals as effective means for connecting teacher evaluation and student learning. Another principal related being able to have more “fierce conversations” [a reference to Susan Scott’s 2004 book of the same name] to evoke change because of evidence and artifacts collected on the standards and indicators. One other principal noted being challenged to coach more using the rubrics.
Finally, Glimpses of Resolution included reports of developing positive relationships with teachers. Universally, principals reported that teachers needed a lot of support in this new evaluation process and expressed a great deal of empathy with their staff as they worked through this process. Principals noted that the process took more time and effort on the part of teachers as well as administrators. Several pointed out ways in which they were trying to lessen the impact on teachers and teaching. One principal described how she was trying to lessen anxiety by alleviating any mystery around the evaluation. She remarked, “So, what I don't want people to do is have any surprises when it comes April and here's your evaluation. I think that will help maybe calm those nerves and alleviate that a little bit more… I feel like teachers do have a tough job.” Another principal described front-loading the idea of a new system:

From the beginning, I told the staff. I said, here's my approach to all of this. We can expend all of our energy fighting it, going against it, whatever you want to say, or we can spend our energy and accept it and move forward….I don't dodge accountability, though. That is never going to happen. If we work together, we can make it manageable.

A few principals referenced bringing a sense of coherence to the bigger picture by trying to align or integrate initiatives. In two cases, principals had teachers examine school or classroom data as a staff or grade level. Teachers then used that data to create student learning goals. Another principal encouraged teachers to work towards group instructional goals as appropriate; for example, an entire school chose to focus on instructional strategies for metacognition.
Many principals pointed out that teachers want to do well and that principals need to help them realize that the evaluation is not about the exact number they receive, it is not about getting 100%. They explained the principal’s responsibility to help teachers change their mindset from “I must get highly effective and a 100% score” to a mindset of improvement: What does this evaluation reveal to me about areas of improvement? What are my next steps? One principal specified, “We are on the side of teachers and we all want to be fair to them and give them the best feedback possible, and with that, it all comes down to not just saying "Good job" but "Good job because--" and "You need to improve because--."

All of the principals expressed that developing positive relationships and trust with teachers was crucial to being taken seriously when it came to feedback. Most reported consciously working on trust as an important part of the process. One principal gave the following picture:

I think with the way I've always approach evaluations is that it has to be a dialogue between the administrator and the teacher, and it can never be felt by the teacher that they're being setup for anything. But it's a very transparent dialogue. If I've missed something, then I need to go back and ask that teacher, "Did I miss this or was it not there and can I see it?" I think building that trust is really important between the administrator and the teacher.

Some principals warned that they should not let the tool or compliance requirements get in the way of the ultimate goal of teacher evaluation. One principal advised the following:
I think what would be my most valuable advice to anybody is, don't let the tool get in the way of what should be taking place in that classroom, that this is about students and their learning. Even though we have to jump through hoops sometimes, and then we have to be compliant sometimes, and then we have to shift and keep it a growth mindset, and we have all kinds of requirements for all different types of reasons. The number one reason why any of us are here is to help those students learn and that every student deserves that opportunity and it's up to us, that's the main responsibility of us as educators is to provide that in a safe and nurturing environment. That can be very difficult to do if you're stressed out about 10 other things that are out of your control. Keeping an eye on the prize, if you will, with student learning, keeping the focus there - and anticipating, or maybe accepting, that it's never going to be perfect and we're going to continue to work through those challenges, but we're here to support, too, and trust is critical.

Finally, one principal summed up the idea that in this cumbersome system there are positive aspects and that, given time, principals would be able to make the most of it:

It’s encouraging to think, “Okay, give me another year or two and I think we’ll have this figured out.” It’s like anything else. Like we said before…one thing you can count on is change and that’s what happens in the system. You just adapt and try to make the best of it. I’m encouraged. I think it will be good.

Summary

Upon analyzing the 16 principal interviews, three major themes emerged relative to principals’ mental models around the new teacher evaluation system. Each theme contained specific subthemes that were explored in Chapter Four. The themes and
subthemes were 1) Tension: The Importance of Doing Teacher Evaluation and Feeling Overloaded and Overwhelmed, 2) Desperately Trying to Figure It Out: It’s a Catch-22 System and Getting a Handle on Time Management, and 3) Glimpses of Resolution. In Chapter Five, the research questions for this study are revisited in light of the data and current research. Implications for practice and future research are suggested.
Chapter Five: Findings and Recommendations

The purpose of this study was to develop an understanding of the mental models held by school principals regarding the implementation of a complex new system of teacher evaluation. The research questions that guided the inquiry were:

1) What mental models (belief systems, assumptions, attitudes, and knowledge) do principals hold about the role of teacher evaluation as a vehicle for change in schools; and,

2) Do different school contexts (elementary, middle, and high school) influence the mental models that site administrators hold relative to teacher evaluation?

A qualitative methodology was used to explore the research questions through the interviews of sixteen K-12 principals. The constructivist grounded theory approach was used to analyze the data collected during the interviews. This final chapter presents the major findings of the inquiry in light of extant literature. Implications for practice will be discussed and limitations of the study will be considered. Finally, suggestions for possible future research will be presented.

Discussion

Results of the analysis of the interviews of the 16 principals revealed an overarching mental model that appeared to reflect a continuum of responses to the new evaluation system. Figure 9 provides a visual construct of the relationship between the findings. The first finding is represented by the notion of Tension. All of the principals indicated that teacher evaluations are important and yet they felt overloaded and overwhelmed by the new system. The Tension created by the conflict between wanting to conduct quality evaluations and feeling overloaded and overwhelmed led to the next step
on the continuum, Desperately Trying to Figure It Out. Though this model was shared by all the principals, it was found that there was a great deal of variability in how they were Figuring It Out. One finding related to this portion of the model indicated that principals viewed the system as a Catch-22; in other words, in trying to conduct the system as intended, it became difficult to complete quality evaluations. Another finding was that principals used a variety of actions and strategies related to Getting a Handle on Time Management. Interestingly, there were a few Glimpses of Resolution as a small number of administrators reported feeling that they were tipping the scales in the direction of finding a reasonable balance between requirements of the system and quality evaluations.

![Figure 9](image)

*Figure 9. Continuum of Major Themes Identified in the Principal Interview Data Regarding Teacher Evaluation. The above graphic represents the major themes and subthemes that emerged from the analysis of the interview data of 16 site principals in Nevada. Tension represents the conflict that occurred between principals’ belief that teacher evaluation is important and feeling overloaded and overwhelmed by the new system. This Tension led to principals Desperately Trying to Figure Out the new system. Principals indicated a Catch-22 system where the time conducting the process actually interfered with quality evaluation. Though the principals shared a mental model, there was great variability in actions and strategies in Time Management. Finally, a few principals indicated some Glimpses of Resolution in synthesizing the elements of the new system.*
Specific differences in principals’ reporting according to school context were not apparent. There were two elementary stand-alone principals, with no assistant principal or dean, who described the same mental model as the whole group. However, these principals indicated they were following the same evaluation procedures, but for many more teachers (up to 34) and without help, so their experience was intensified.

The most notable finding was that all of the 16 principals interviewed held the same general beliefs and attitudes about teacher evaluation. Principals expressed the belief that teacher evaluation is important and that the ultimate purpose is to increase teacher capacity to increase student learning. Nuances included 1) believing that the process was valuable, though the writing and delivery of the final evaluation did not necessarily produce change; 2) providing appropriate and timely feedback was important; and 3) building relationships and trust were crucial to effective conversations with teachers around performance. This last nuance included exhibiting understanding and empathy for teachers with respect to the challenges of the new evaluation expectations.

According to Hirsch, Psencik, and Brown (2014),

Leaders in learning systems inspire confidence in those around them. They gain others’ trust by acting with sincerity, reliability, and competence. Those who trust the leader believe the leader has their best interest at heart, as well as the organization’s mission. (p. 92)

For the principals, building trust was a mechanism for building learning in the school.

There is consensus in the literature that learning-focused conversations and quality feedback for teachers can increase teacher effectiveness and lead to student learning (Blase & Blase, 2004; Danielson & McGreal, 2000; Lipton & Wellman, 2013;
Marshall, 2009; Marzano, Frontier, & Livingston, 2011; McGreal, 1983; Platt & Tripp, 2014). The data from this study suggested that principals’ perceptions of successful evaluation practices were in line with the research. Blase and Blase (2004) found that one of the three primary elements of successful instructional leadership was conducting instructional conferences. These conferences included making suggestions, giving feedback, modeling, using inquiry, and soliciting advice and opinions from teachers. These descriptors align with what principals were experiencing in the schools.

Descriptions indicated that principals viewed evaluation through two lenses. The first lens was supervision and coaching, that is, the process. Supervision and coaching was where the principals felt they had the most impact on increasing teacher capacity and where they wanted to spend the bulk of their time. The second lens, that of evaluation, was considered more pro forma and was described as less impactful on teacher performance, and yet it took up a great deal of time. Marshall (2009) discovered that principals spend approximately 99.9 percent of their time on activities other than being in classrooms and conceded that if all the steps to conventional supervision and evaluation “were implemented skillfully, supervision and evaluation would be a significant force for improvement” (p. 21). The reality described by the principals is in line with the research that shows that leaders are pressed to find ways to balance their time effectively. Marshall (2005) recommended to “use short observation visits to write teachers’ final evaluations. Dispensing with elaborate, announced evaluations is a huge time-saver, and once a trusting climate has been established, it’s the ideal scenario” (p. 734).

Principals struggled with internalizing the rubrics and managing the requirements of the new evaluation system. The new teacher evaluation was just one of the many
initiatives under their purview and they were feeling overloaded and overwhelmed. Even though the evaluation of teachers had always been part of an administrator’s duties, the time, paperwork, and cumbersome new system created additional challenges. These challenges combined with the desire to be efficient and accurate in a high-stakes system caused strong feelings of Tension among the administrators.

The Tension created by the principals’ strong feelings of responsibility toward their teachers and students and the equally strong feeling of being overloaded and overwhelmed was evident in their descriptions. Serendipitously, these feelings of Tension were the impetus for a certain amount of exposing and sharing of the mental model among administrators. According to Senge (2006),

Though highly personal in nature at one level, effective work with mental models is also pragmatic, that is, it is tied to bringing key assumptions about important business issues to the surface. This is vital because the most crucial mental models in any organization are those shared by key decision makers. (p. 176)

Almost all of the principals described sharing information with colleagues to some degree to deepen their knowledge about the new evaluation procedures and standards. Half of the principals described working as an administrative team on site to exchange ideas for providing feedback or writing the evaluation. Some principals met as a district administrative team and others collaborated in small groups with principals from other schools, visiting classrooms and calibrating scoring of observations. Whether planned or spontaneous, the fact that administrators met to air their frustrations, discuss and share learning, or brainstorm strategies for accomplishing the task of evaluation provided a type of unconscious double-loop learning opportunity. This collaboration was
a positive byproduct, leaning toward the unearthing and sharing of principals’ mental models and is supported by the literature.

As Senge (2006) pointed out, the purposeful sharing of mental models among key decision makers is important for the successful implementation of initiatives as well as for increasing the level of learning of the organization. Elmore (2000) declared, “Improvement occurs through organized social learning, not through the idiosyncratic experimentation and discovery of variously talented individuals” (p. 25). Developing and supporting organized social learning among administrators in a planned and purposeful (rather than serendipitous) manner would be beneficial for calibrating mental models around teacher evaluation as the new processes continue to be refined. Principals engaging with other principals in dialogue around their own learning and evaluation processes could provide a double-loop feedback opportunity that would deepen the chances of success for the initiative. Fullan (2001) referred to this as “learning in context over time” (p. 125) and described the benefits as follows:

This is fantastic insight: learning in the setting where you work or learning in context, is the learning with the greatest payoff because it is more specific (customized to the situation) and because it is social (involves the group). Learning in context is developing leadership and improving the organization as you go. Such learning changes the individual and the context simultaneously. (p. 126)

Some examples of learning in context include intervisitation, monthly principal support groups, principal peer coaching, supervisory walkthrough, district institutes, principals’ study groups, or individualized coaching (Fullan, 2001).
The second finding revealed that, although there was a shared mental model around the importance of evaluating teachers, there was variability in interpretation and implementation of the NEPF. Principals were Desperately Trying to Figure Out how to manage the entire system: learning the rubrics, identifying the indicators in action, and refining the process within a doable, yet effective, time commitment.

Because the new NEPF evaluation system was a legislative mandate, it was generally viewed as an externally imposed initiative. Interviewees described a conflict between implementing the system and maintaining the integrity of the purpose of the system. This conflict is represented in the literature by the concepts of external and internal commitment that evolve during the change process. Argyris (2000) described external and internal commitment as follows:

Both are valuable. Both can lead to persistence, endurance, and vigilance. Both can coexist. But the consequences to which each leads are naturally contradictory. Commitment can be external or internal. These differ in how they are activated and in the source of energy they utilize. External commitment is triggered by management policies and practices that enable employees to accomplish their tasks. Internal commitment derives from energies internal to human beings that are activated because getting a job done is intrinsically rewarding. (p. 40)

Though the principals espoused a commitment to the ultimate purpose of teacher evaluation (increasing the capacity of teachers to increase student learning), there was evidence of different levels of commitment by administrators to the initiative. Principals exhibited external commitment through accomplishing the practical tasks such as scheduling, meeting deadlines, creating walkthrough tools, using feedback forms, etc.; in
other words, the management practices and policies that helped them accomplish the task.

At the same time, principals displayed, to varying degrees, a focus on internal commitment. Internal commitment was displayed through deepening individual learning around the evaluation indicators, scaffolding teacher leadership, providing meaningful feedback, engaging teachers as learning partners, placing emphasis on goal setting, etc., that is, focusing on actions to advance internal commitment to the initiative. Both external and internal types of commitment were identified among all of the principals and each principal displayed a certain level of both.

The descriptions provided by the principals regarding multiple initiatives and the time required to complete evaluations indicated that principals are indeed good soldiers who will do whatever it takes to get the job done. However, getting the job done from the external commitment point of view could lead to successfully completing a compliance system. The theme of a Catch-22 System revealed that compliance might actually interfere with the ultimate goal of teacher evaluation: increasing teacher capacity to increase student learning. Ironically, according to most principals, fulfilling the requirements of the system took time away from the important parts of the system: providing feedback and support where needed.

Bateson (1972) described these types of no-win situations as a double bind. According to Bateson’s research on schizophrenia, six conditions must be in place for a double bind to occur: 1) two or more persons (one of which is the “victim”), 2) repeated experience (a recurrent theme in the experience of the victim), 3) a primary negative junction (some form of punishment – Do not do this and I will punish you or if you do this, I will punish you), 4) a secondary negative junction conflicting with the first
(commonly communicated by the victim), 5) a tertiary negative junction (something that prohibits the victim from escaping the situation), 6) the complete set of ingredients is no longer necessary (occurs when the double bind pattern takes hold). Bateson explained that any individual could break down under these repeated conditions and respond in a defensive manner. The defense response is a resolution and could be productive such as identifying a radical solution to the problem. More likely, the resolution would be detrimental, such as developing evasive or unacceptable coping behaviors.

Certainly, this description is a simplified explanation of an extremely complicated concept, but in terms of the Catch-22 System theme for the school principals, it could provide some insight into the various actions and reactions (resolutions) they were experiencing. As they were Desperately Trying to Figure Out this new and cumbersome system, the principals could have been managing a double bind situation.

The theme of Getting a Handle on Time Management illustrated how complex the role of the principal is in today’s educational environment, especially during times of change. Gronn (2003) described the intensification of work for school leaders in the twenty-first century as a real and problematic situation, which has now become the norm rather than the unusual. Gronn (2003) noted,

> With intensification under NPM [new public management], however, school leaders’ role demands have become numerically large and exceedingly complex, and the constraints they face extensive and imposing, with the result that in many instances the opportunities for widespread influence and transformative agency (as enshrined in standards statements) have been minimized… (p. 84)
Gronn’s observations were consistent with principals who described their role as becoming so complex that it was hardly doable.

This was significant because it has implications for future results of the new evaluation initiative. Principals and their actions seemed to fall on a continuum of change called the Implementation Dip. Hall and Hord (2001) defined the Implementation Dip as follows:

This is the period where a change has been designed and developed by implementers at the school or district level and introduced to colleagues, or perhaps has been transferred from a vendor or another setting and introduced. As individuals struggle to make the change ‘work,’ they go through the valley, or dip, of difficulties before they reach the top and emerge at a higher level, which is an improved status. (p. 192)

Principals opting for only external types of solutions may end up with accomplishing minimal compliance. However, principals committing to the internal moral purpose of evaluation may be able to persevere through the Implementation Dip and achieve higher status on the other side, possibly seeing changes in teaching and learning.

Many principals did not exhibit full confidence in the overall effectiveness of the evaluation system to bring about significant change. A push and pull situation seemed to exist between developing a compliance system and a capacity-building system. Principals were skeptical about being able to manage the requirements of the system as currently expected and the ability to provide meaningful feedback, coaching, and other support for teachers that would result in a capacity-building system that actualized long-term changes in teaching and learning.
A final finding was indicated by a small group of participants who were experiencing Glimpses of Resolution. As a bookend to the double bind, these principals were finding ways of both managing the external commitment types of processes in the evaluation system and maintaining focus on the purpose of evaluation: increasing the capacity of teachers to increase student learning.

According to Fullan (2001), “Leadership, to be effective, must spread throughout the organization” (p. 57). Fullan advised that site leaders must develop leadership capacity in others to spread and sustain a focus on continuous improvement. Some principals were encouraging teachers to take on leadership roles, including them as learning partners, or providing time for them to support and mentor colleagues. Other administrators reported noticing changes in conversations about instruction as teachers deepened their understanding about the evaluation indicators. Some reported a shift from single events of effective instruction (the dog and pony show) to more embedded, long-term use of effective instructional strategies. A few principals mentioned increasing their own capacity to have more effective conversations with teachers through the use of the rubric descriptors.

In sum, the major findings in this study revealed that principals held a shared mental model regarding teacher evaluation. Through the analysis of the descriptions given by the principals of how they were trying to figure out the new evaluation system, it was revealed that the role of school leader is extremely complex in times of educational change. Research indicates that approaching complex issues from an external commitment viewpoint, that is, from a practical, transactional perspective, could lead to a compliance oriented solution. On the other hand, approaching challenges from an internal
commitment viewpoint, that is, focusing on the moral purpose of the initiative, could lead to increased learning and positive change (Fullan, 2001). The graphic below, Figure 10, provides a visual representation of the connections between the findings and the research.

Figure 10. Relationship between Major Findings. The above graphic represents the relationship between the major findings of the data analysis and research. Principals shared a mental model: Because teacher evaluation is so important and the process is one more thing on my already overloaded plate, I am feeling overwhelmed. I am trying desperately to figure this out. As principals were desperately trying to figure out the new evaluation system and trying a myriad of approaches, the complex role that school leaders play in the change process and in teacher evaluation was confirmed. Principals opting for only external types of solutions (practical) may end with accomplishing compliance. Principals committing to the internal main purpose of evaluation (intrinsic) may be able to persevere through the Implementation Dip and achieve higher status on the other side, possibly seeing changes in teaching and learning.

Implications for Practice

The findings of this study indicate that the school leaders represented were deep into a change process with regard to the evaluation of teachers. Indeed, the descriptions supplied by the 16 principals regarding their experiences with teacher evaluation illustrate how complex the role of principal has become. As this initiative continues to
roll out and be refined over time, and acknowledging that other initiatives will continue
to be added, the implications for practice must focus on elements that provide support for
principals.

One such support for principals might be the flexibility to streamline the
evaluation process. For instance, teachers who have demonstrated excellence in teaching
may not require evaluations every year. Incorporating a rotation cycle of evaluation for
teachers who exhibit effective or highly effective ratings over time would ease the burden
on principals and allow them to focus on teachers who needed their support. As another
example, utilizing a system of peer evaluation would provide content expertise to
teachers and principals as well as lessen the time required of principals for observation
and feedback. An added benefit of this approach would be increasing the capacity of
teacher leaders in the school building and across the district.

In terms of professional learning support for administrators, a guided, ongoing
study of the evaluation standards and indicators in a collaborative setting would benefit
principals in several ways. First, they would gain deeper knowledge to define the
concepts embedded in the indicators that were the most problematic for themselves and
for the teachers. This understanding would increase principals’ observation skills and
allow them to provide concrete supports and resources for teachers in confusing areas of
the rubrics. Developing a deeper common understanding of the indicators would increase
consistency in both written and verbal feedback as well as in the final written evaluations.
The collaborative aspect of this learning would increase shared knowledge within schools
and across districts. Finally, additional opportunities to increase expertise in coaching and
feedback practices would further empower administrators to support teachers as needed.

**Limitations of the Study**

Limitations inherent to this study are as follow. First, the sample size was small, including only 16 site principals from four small to mid-sized school districts. High schools and middle schools were represented by only four principals each. Elementary schools, the largest number of schools in the region, were represented by only eight principals. The information provided by the principals refers only to the perceptions of administrators in four school districts, though the teacher evaluation initiative is in place in sixteen of the seventeen school districts in the state. The sample size was too small for conclusions to be extrapolated to other contexts, though some of the discussions might apply to specific situations elsewhere.

Second, only one type of data was examined, the interview. Though there was confirmation across the data on major themes, including other types of artifacts or data collection might have been confirmatory, produced other results, or revealed gaps in the themes.

Additionally, the information collected through the interview was recorded and completed in a face to face situation. Recording could have been a factor in how each principal described his or her experiences. This type of non-anonymous, more personalized situation could have influenced interviewees to hold back information regarding certain subjects. The self-reporting aspect might have produced different results than a more neutral type of data collection method.

Finally, the focus of this study was intended to be centered on teacher evaluation in general. Because the timing of the interviews coincided with the advent of the new
evaluation system introduced in Nevada, the focus quickly became centered on the new system. This narrowing of focus could have influenced the findings and implications for future research.

**Future Research**

There are several implications for future research as a consequence of this study. First, additional research regarding the mental models of principals and school change would add to the growing body of information in that area. Next, according to the principals in this study, a great amount of time and resources have been focused on teacher evaluation. A follow-up inquiry would be interesting to determine whether this focus is making a difference in increasing the capacity of teachers to increase student learning. In terms of leadership, it would also be interesting to discover which leadership actions are having the greatest affect on teachers and student learning in the context of this new system. A study of the principals’ own evaluation process could also be revealing in terms of the overall evaluation picture.

Research regarding to what extent teaching behaviors might be changing would be a logical extension of this study and would provide additional information as to the effects of the system on classroom instruction. Finally, additional research on the use of teacher evaluation as a vehicle for change in schools would be of importance as districts and states across the nation implement this theory. State data are beginning to be collected that could be analyzed for information purposes in this state as well as for comparison with other states.
Conclusion

In conclusion, the implementation of teacher evaluation as a leverage point for increasing the effectiveness of instruction in schools has taken hold across the nation. Findings of this study indicated that the role of the school principal has become increasingly complex and that principals’ mental models play a significant role in the change process as related to teacher evaluation.

These findings have implications for how principals are supported in their role as school leader with respect to the social, economic, political, and commercial forces that are driving educational changes in the United States. First, pre-service administrator development programs must align with these forces and be supported by research which will enable future school leaders to have greater probabilities of success. In addition, universities could engage in inservice activities to provide ongoing extended learning opportunities to support the success of practicing school administrators.

Next, state and district level decision makers must be aware of the repercussions of mandating complicated multiple, overlapping, and/or conflicting initiatives with minimal support for school leaders. This study of the efforts of school principals to implement a complex new teacher evaluation system illustrates some of the various issues that can develop. Even so, all principals in the study have worked to implement this initiative with fidelity.

Finally, professional learning providers such as district personnel, universities, or state and regional entities must be sensitive to the issues and needs of school administrators. These organizations must develop an understanding of the forces that drive change in public education and should work in collaboration with local education
agencies, school districts, and administrators to provide appropriate support. Specifically, these entities should consider the stress that significant changes cause for principals who must implement and manage these changes.

As Senge (2010) averred, “Where a whole system approach has been taken seriously over the past decade, there have been significant improvements in student achievement” (p. viii). Schools, districts, and states that have fostered the concepts of becoming a learning organization have focused on increasing the capacity of individuals as well as advancing the collective efficacy of educators through a systems approach. According to Senge (2006) and Argyris and Schön (1996), to move systems forward, mental models must be unearthed and examined or they become barriers to success and actually prevent organizational learning. Understanding mental models of educational stakeholders is critical to enable schools to become learning organizations. If schools are to shift from being compliance-oriented institutions to learning organizations, the influence of mental models must be understood or the current paradigms of stakeholders will prevail.

This study indicated that, according to the principals, the purpose of teacher evaluation is to increase teacher capacity to increase student learning. In essence, this mental model can be linked to Senge’s (1990) five disciplines; thus, it can be linked to learning organizations. However, for schools to become learning organizations, principals must have sufficient support; without such support the risk exists that they will lead organizations that exhibit the characteristics of the organizational learning disabilities. In summary, the overall conclusion is that principals appear to be very willing to implement and lead complex initiatives; however, they will need meaningful, aligned support.
References


National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest website: http://ies.ed.gov/ncee/edlabs


http://nspf.doe.nv.gov/Home/NVReportCards

http://nspf.doe.nv.gov/Home/FAQ
Nevada Department of Education (August 2015). *Six things you need to know about school ratings from the 2014-2015 school year.*

http://nspf.doe.nv.gov/Content/PDF/six%20things.pdf


WestEd. (2009). Western Nevada regional training program T4S program. Carson City, NV: WNRTCP
Appendix A

Sample Recruitment Protocol

Each participant will be initially contacted by phone by the researcher. The following phone protocol will be used:

“Hello, this is Kirsten Gleissner, Director of __________. I am also a doctoral student at the University of Nevada, Reno. I am conducting a study of the beliefs, assumptions, attitudes, and knowledge of school principals regarding teacher evaluation. I would love to have your insights regarding the evaluation process for the study.

If you choose to participate, I will ask for about an hour of your time to conduct an in-depth interview with regard to your thoughts about the teacher evaluation process. The interview will be audio recorded so that I may give my full attention to our conversation and so your thoughts are accurately represented. Recording the interview will also allow me to go back to our conversation at a later date. After I transcribe the interview tapes, I will send you a copy so you can read the transcript to make sure that your thoughts were accurately recorded.

The interview may occur at a time and place of your choosing. Your identity and statements will remain confidential and you will not be identified in the write-up of the study. This study has been approved by both the university and the District.

Would you be able to participate in this study?”

- If the participant answers yes, an appointment will be made for the interview.
- If the participant answers no, the participant will be thanked for his or her time.
## Appendix B

### Principal Interview Questions

<table>
<thead>
<tr>
<th>Phase</th>
<th>Questions</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial Open-ended Questions:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General background</td>
<td>1. Would you describe your current role and how you came to be here?</td>
<td>Ice breaker</td>
</tr>
<tr>
<td></td>
<td><em>Looking for:</em></td>
<td>Knowledge</td>
</tr>
<tr>
<td></td>
<td>• Current school, position, goals of the school</td>
<td>Experience</td>
</tr>
<tr>
<td></td>
<td>• Previous schools, positions (teaching, admin, counseling, coaching, etc.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Previous districts, states (systems)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• How long in education total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Other possible experiences (coming from industry?)</td>
<td></td>
</tr>
<tr>
<td>Specific background</td>
<td>2. When did you first begin doing teacher evaluations?</td>
<td>Knowledge</td>
</tr>
<tr>
<td></td>
<td>• What was it like then?</td>
<td>Experience</td>
</tr>
<tr>
<td></td>
<td>• What were you thinking about your experience during that time?</td>
<td>Attitudes</td>
</tr>
<tr>
<td></td>
<td>• Who, if anyone, influenced the way you conducted your evaluations?</td>
<td>Beliefs</td>
</tr>
<tr>
<td></td>
<td>Tell me about how s/he influenced you?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• How would you describe the type of administrator you were then?</td>
<td></td>
</tr>
<tr>
<td><strong>Intermediate Questions:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEPF System</td>
<td>3. When did you first hear about the NEPF?</td>
<td>Knowledge</td>
</tr>
<tr>
<td></td>
<td>• Could you tell me about your initial thoughts and feelings when you learned about it?</td>
<td>Experience</td>
</tr>
<tr>
<td></td>
<td>• How did you learn to use the current system?</td>
<td>Assumptions</td>
</tr>
<tr>
<td></td>
<td>• Can you tell me about any changes you have made in the way you do teacher evaluations now?</td>
<td>Attitudes</td>
</tr>
<tr>
<td></td>
<td>• What most contributed to these changes?</td>
<td></td>
</tr>
<tr>
<td>Experiences</td>
<td>4. What are some experiences that stand out to you from doing evaluations this year?</td>
<td>Experience</td>
</tr>
<tr>
<td></td>
<td>• How would you describe how you are feeling about how you are conducting evaluations at this time?</td>
<td>Attitudes</td>
</tr>
<tr>
<td>Phase</td>
<td>Questions</td>
<td>Question Focus</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td></td>
<td>• How do you think your colleagues might describe their experience? Do you think they would say the same things or different things than you?</td>
<td></td>
</tr>
<tr>
<td>Reactions</td>
<td>5. Tell me about some of the reactions you have had around conducting teacher evaluations.</td>
<td>Assumptions Beliefs</td>
</tr>
<tr>
<td></td>
<td>• How have the reactions you have experienced influenced how you are doing teacher evaluations?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Can you tell me about how you use the information you gain from teacher evaluations?</td>
<td></td>
</tr>
<tr>
<td>Purpose</td>
<td>6. What do you think are the main purposes of a teacher evaluation process?</td>
<td>Beliefs Assumptions Attitudes</td>
</tr>
<tr>
<td></td>
<td>• Can you tell me about how this evaluation process fits with that?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Could you tell me about how your views may have changed around teacher evaluation since you have been in the new system?</td>
<td></td>
</tr>
<tr>
<td>Student connection</td>
<td>7. Tell me about what you see as the relationship between teacher evaluations and student learning?</td>
<td>Beliefs Assumptions Attitudes</td>
</tr>
<tr>
<td>Ending Questions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Future</td>
<td>8. After having gone through this experience, what advice would you give to someone who has just discovered that s/he will be responsible for doing teacher evaluations?</td>
<td>Knowledge Assumptions Beliefs</td>
</tr>
<tr>
<td></td>
<td>• Is there something you know now that you wish you would have known earlier?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9. Is there something that you might not have thought about before that occurred to you during this interview?</td>
<td>Knowledge</td>
</tr>
<tr>
<td></td>
<td>10. Is there something else you think I should know to understand this process better?</td>
<td>Assumptions Beliefs</td>
</tr>
<tr>
<td></td>
<td>11. Is there anything you would like to ask me?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• May I contact you with clarification questions,</td>
<td></td>
</tr>
<tr>
<td>Phase</td>
<td>Questions</td>
<td>Question Focus</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td>if necessary? (What would be best? Email? Phone?)</td>
<td></td>
</tr>
</tbody>
</table>

**Reserve Questions**

**Optional:**

- Tell me about how you have grown as a school leader through this process.
- Tell me about the strengths that you discovered or developed through this process.
- What do you most value about yourself now?
- What do others most value in you?
Appendix C

Consent Form

Consent Information for the Study of Mental Models Held by School Principals Regarding Teacher Evaluation

I am conducting a research study to learn about the mental models held by school principals regarding teacher evaluation.

If you volunteer to be in this study, you will participate in a one-to-one interview where you will be asked some questions in a conversational way about your knowledge, assumptions, beliefs, and attitudes regarding teacher evaluation. The interview will be audio-recorded to ensure the accuracy of your comments and to allow focused attention on your responses. The interview will be transcribed and you will be sent a copy to review.

Your participation should take about one to one and a half hours.

This study is considered to be minimal risk of harm. This means the risks of your participation in the research are similar in type or intensity to what you encounter during your daily activities.

Benefits of doing research are not definite; but we hope to learn about what assumptions and beliefs principals hold about the teacher evaluation process, what their attitudes are towards it, and what knowledge they possess. This deeper understanding may lead to better ways of supporting principals in the teacher evaluation process. In addition, information may be gained about providing more effective professional development related to teacher evaluation. There are no direct benefits to you in this study activity.

The researchers and the University of Nevada, Reno will treat your identity and the information collected about you with professional standards of confidentiality and protect it to the extent allowed by law. You will not be personally identified in any reports or publications that may result from this study. The US Department of Health and Human Services, the University of Nevada, Reno Research Integrity Office, and the Institutional Review Board may look at your study records.

Required Language
You may ask questions of the researcher at any time by calling Kirsten Gleissner at (775) 230-8642 or by sending an email to gleissnerk@yahoo.com.
Your participation in this study is completely voluntary. You may stop at any time. Declining to participate or stopping your participation will not have any negative effects on your current situation or circumstances.

You may ask about your rights as a research participant. If you have questions, concerns, or complaints about this research, you may report them (anonymously if you so choose) by calling the University of Nevada, Reno Research Integrity Office at 775.327.2368.

Thank you for your participation in this study!

Kirsten Gleissner