Interpreting Teacher Evaluation Policies:
The Perspectives of Local and State-Level Policy Actors in Two U.S. States

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Jane F. Gilles

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Karen R. Seashore, Ph.D., Adviser

December 2015

Acknowledgements

Earning a doctoral degree, like many of life's accomplishments, is only possible with the support of others. It is with deep gratitude that I acknowledge those whose guidance and encouragement have meant the most to me as I have traveled this path.

My adviser Dr. Karen Seashore has influenced me in countless ways. In Karen's classes I learned to see myself as part of a community of researchers. In our advising sessions I learned to view my career and academic work from a broader perspective, as part of a regional, national, and international education policy field to which I could contribute. I left every encounter with Karen feeling smarter and more confident than I had felt before. (Truly, I recall telling friends that I could almost feel my I.Q. going up each time Karen and I met!) Karen, I am forever grateful that *you* led me on this journey.

Whenever I reflect on my Ph.D. work, I will think of Dr. Nicola Alexander. Her education politics and policy courses lit a spark in me and laid the groundwork for my academic passions. I have benefitted greatly from Nicola's wisdom and kindness, and I will always credit her for making me a better thinker and a better writer. I feel privileged and honored to have had Drs. Jodi Sandfort and Andrew Furco serve on my dissertation committee. Thoughtful in their critiques and encouraging in their guidance, Jodi and Andrew have both exposed me to new perspectives that have sharpened my skills as a researcher and policy analyst.

The participants in this study gave generously of their time, and I am indebted to them for their willingness to share their thoughts and experiences in support of my dissertation research. Without exception, participants in this study exhibited a deep commitment to faithfully implementing new teacher evaluation policies for the benefit of students. I am grateful to these professionals for their assistance with this study.

I have been fortunate to work with excellent colleagues throughout my career. They have inspired me with their passion for teaching, leadership, and service. We choose this profession because we care about kids, and I have never had to look far to see examples of educators who are willing to work ever harder and sacrifice ever more – always focused on the students we serve. I thank these colleagues for enriching my professional life.

My family and closest friends have exhibited endless patience over the six years since I began the pursuit of my Ph.D. They have listened intently to my reports of both the most exciting and most difficult aspects of my program, shown interest in my research, helped with literally thousands of tasks I couldn't get to during the busiest times, and never, ever complained about hearing yet another "no" in response to an invitation. Their love and support has sustained and humbled me.

Dedication

This dissertation is dedicated to the memory of my parents, John and May Gilles.

They were with me every step of the way.

Abstract

This multi-case study explores how local policy actors – teachers and administrators – interpret new state teacher evaluation policies developed in response to federal pressures. In the study's first phase, local policy actors from six small school districts in two U.S. states were interviewed about their efforts to develop teacher evaluation systems to address new state policies. In the second phase of the study, interviews with state-level policy actors in the two states expanded on themes from the first phase. The study's conceptual framework recognizes policy interpretation as a distinct phase in the policy process, and it addresses the context, influences, and multiple actors involved in policy interpretation. Using an open coding approach consistent with grounded theory, interviews were analyzed, and supporting documents were used to triangulate the data. The study finds that local actors embrace teacher evaluation as a means of professional growth, yet they hold serious concerns about including measures of student growth in teacher ratings. In addition, the study reveals that the density of collaboration at the local level influences the scope of policy interpretation; that local policy actors rely on state-level stakeholders – who are shown to serve as policy intermediaries – to aid in policy interpretation; that the messages of policy intermediaries influence local actors' understanding of policy goals and attitudes toward implementation; that collaboration among policy intermediaries benefits local policy actors; and that state political culture influences policy activity at both the state and local levels. Since this case study was conducted in only two states and only a small number of local and state-level policy

actors were interviewed within each state, a total of 35 participants across both phases of the study, the research is limited and findings cannot be generalized.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

Currently there is unprecedented focus on teacher evaluation in the U.S. as well as in other countries. Within the U.S., at both state and federal levels, teacher evaluation has risen in prominence as a key approach to improving the quality of teaching and learning and as a tool for increasing student academic performance. As a consequence of emerging state and federal debates, policy actors are operating in an environment that presumes that more and better evaluations of teachers is one of the solutions to a perceived crisis in American education. Because the U.S. constitution delegates responsibility for education to the states, many versions of enhanced teacher evaluation policies have been adopted across the country, and most states also allow local school districts some degree of autonomy in the design of teacher evaluation systems.

The literature indicates that state and federal attention to teacher evaluation has grown dramatically over the past thirty years, with a sharp increase in state policy development since President Obama took office in 2009 and his administration initiated multiple federal policies encouraging changes to the practice of evaluating both teachers and school principals. While only four states had policies on the books that required evaluating or assisting in-service teachers prior to 1970 (Tracy & Smeaton, 1993), currently, all 50 U.S. states and the District of Columbia have policies requiring teacher evaluation, and all but a handful of those with recently adopted policies had implemented them by 2015 (Center on Great Teachers & Leaders, 2013; Hull, 2013; Doherty & Jacobs, 2013; National Council on Teacher Quality, 2014b; Doore, B., Jorgensen, Doore,

S. & Mason, 2013). In response to the policy activity at the federal and state levels, the last few years have found policy actors in local school districts developing and implementing new evaluation systems all across the U.S.

The increase in teacher evaluation policy activity has spurred controversy in states and school districts, and teacher evaluation makes headlines on a regular basis. Indeed, a LexisNexis® news search on the term "teacher evaluation" for the first six months of 2015 yielded over 1,750 results.[1] Following are some high-profile controversies that have occurred in recent years:

- The publication of teacher evaluation ratings in New York City (where the *New York Post* identified the city's "worst teacher" by name) and Los Angeles (where the family of a veteran teacher blamed his suicide on the fact that his "less effective" rating was made public by the *Los Angeles Times*) (Pathe & Choe, 2013).
- Ongoing controversies in New York State, with Governor Cuomo and teachers unions repeatedly at odds over issues such as the public release of teacher ratings (as mentioned above), the weight given to student test results in evaluation systems, the ability of teachers to appeal evaluation ratings, and changes to state teacher evaluation requirements (Sawchuk, 2013; Liu, 2015).
- Lawsuits in Tennessee and Florida over policies that evaluate teachers based in part on test scores of students they don't teach; lawsuits in Houston and Syracuse

[1] LexisNexis is an online research database providing access to a comprehensive set of news sources, including local, regional, national, and international newspapers (LexisNexis, 2015).

charging that the value-added models[2] used in local teacher evaluation systems are arbitrary and unreliable; a lawsuit in New Mexico over errors in teacher evaluation results purportedly caused by rushed implementation and faulty data; and a lawsuit against 13 California school districts claiming that agreements with local teachers unions prohibiting the use of standardized test scores in evaluating teachers violate state law (Brown, 2015; Shimura, 2015).

- Teacher strikes in Chicago and Seattle in which teacher evaluation policies were a key issue, with Chicago teachers objecting that the weighting of test scores exceeded the state mandate, and Seattle teachers seeking to remove standardized test scores as a teacher evaluation measure; in both cases, the teachers prevailed (Pearson, 2012; Higgins, 2015; Sawchuk, 2015).

Some scholars assert that the increasing attention to teacher evaluation is part of a teacher accountability movement that seeks to reform public education by applying market-based values (competition, extrinsic rewards, quantifiable results) to the profession of teaching (Ravitch, 2012; Darling-Hammond, 2014; Connell, 2013; Piro & Mullen, 2013; Murphy, Hallinger & Heck, 2013). Some charge that proponents of teacher accountability reforms perpetuate faulty assumptions about teaching and learning, namely that standardized tests are effective measures of student learning, that students can accurately demonstrate their learning by means of a standardized test, and that successful teaching is evidenced by increases in student test scores (Kumashiro, 2012). Further, scholars who hold a critical lens to the teacher accountability movement assert that

[2] Value-added modeling is a statistical method that measures a teacher's contribution to the growth in the achievement of students. (Value-added modeling is addressed in Chapter Two.)

reformers place the blame for a supposed crisis in American schooling squarely on the shoulders of teachers (Taubman, 2012; Kumashiro & Meiners, 2012; Shiller, 2015; Stern, 2013). In *Bad Teacher! How Blaming Teachers Distorts the Bigger Picture*, Kevin Kumashiro explains that the narrative behind the teacher accountability movement ignores systemic challenges, narrowing the focus to the classroom level. He explains that the logic of this narrative goes as follows: "Good teachers make for good schools, and since we hear repeatedly that our schools are bad, so too must be our teachers. At least some of them" (2012, pp. 20-21).

With the upsurge in teacher evaluation policy activity in recent years, the intense controversies surrounding the new policies, and concerns about a market-oriented teacher accountability movement, a study investigating teacher evaluation is of significant interest at this time, and responds to a need to understand how national and state pressures are affecting policy actors. The remainder of this introductory chapter articulates the study's research purpose and questions, introduces the conceptual framework for the study, outlines limitations and delimitations, defines important terminology, and provides an overview of the remaining chapters.

Research Purpose and Questions

The purpose of this research was to explore the efforts of local policy actors as they interpreted new teacher evaluation policies and developed local teacher evaluation systems in response to external policy mandates. This study employed case study methodology to address the main research question: How do local policy actors interpret teacher evaluation policies? Sub-questions included:

1. How do local policy actors interpret the *meaning* of teacher evaluation policies?
2. How do local policy actors experience the *process* of interpreting teacher evaluation policies?
3. How do state-level policy actors influence local responses to teacher evaluation policies?

The central research question can be addressed from two perspectives: *content* and *activity*, as indicated by the first two sub-questions. From the first perspective, the research question asks about the interpretations of local policy actors regarding the *content* of new teacher evaluation policies. This perspective concerns the *meaning* of the policy. From the second perspective, the research question asks about the *process* of interpreting the new policy. This activity perspective concerns the *actions and experiences* of local policy actors. The third sub-question addresses the influence of state-level policy actors on local policy activity as local policy actors interpret teacher evaluation policies and plan for their implementation.

Introduction to Conceptual Framework

The conceptual framework for this study, which is articulated fully in Chapter Two, recognizes policy interpretation as a separate phase in the policy process. The framework addresses the context, influences, and actors involved in policy interpretation, incorporating these concepts: the ecology metaphor for educational policy articulated by Weaver-Hightower (2008); Elazar's political culture typology (1970); and the conception of policy intermediaries as aids in the policy process, which has been articulated by

Honig (2004), Hill (2003), and Hamann & Lane (2004) among others. This conceptual framework centers on local policy actors as learners who influence and are influenced by state policy intermediaries. In this framework, both sets of policy actors reside within and are affected by state political culture and the broader policy ecology.

Policy ecology. The literature related to political culture and policy intermediaries is addressed in Chapter Two, but a brief description of Weaver-Hightower's (2008) ecology metaphor for educational policy is included here to help frame the study's context. The policy ecology metaphor offered by Weaver-Hightower is an extension of Firestone's (1989) ecology of games, in which various social games – such as the legislative game, the research game, and the classroom game – operate together within an ecological system. Weaver rejects Firestone's notion of games, yet he builds on Firestone's metaphor, illustrating how a complex ecological system both affects and is affected by a particular policy. In Weaver-Hightower's metaphor, as in science, a system of relationships exists in which each factor in the environment influences the others, and "many complex interrelationships among them are required to sustain the system" (p. 155). For Weaver-Hightower, a policy ecology consists of multiple actors, including but not limited to governments, policymakers, voters, the media, educators, students, parents, advocates, and the organizations to which these actors belong. The ecology metaphor supports the idea that policy can be viewed both as *text* – a tangible document that can be read, interpreted, and analyzed – and as *discourse*, controlling what is thought in regards to the policy. In addition, Weaver-Hightower's conception of policy as ecology illustrates the importance of the various types of relationships among actors and acknowledges the

limits and opportunities offered by the structures and processes in which policies are formulated and enacted.

Study Limitations and Delimitations

This study has certain limitations and delimitations. First, the study has a qualitative design, specifically employing case study research; thus, this research is not generalizable to the broader population of U.S. states and school districts, but instead is intended to inform future research and theory building. Next, since the primary data collection method utilized in the study was the telephone interview, the study is limited by my lack of opportunity to observe non-verbal cues displayed by participants during interviews. Regarding the selection of participants, as described in Chapter Three, I relied on the school superintendent in each participating district to honestly and accurately recommend Phase I participants who met the criteria for participation. In the study's second phase, a purposive sampling approach was utilized, which may have restricted the perspectives that were explored and limited the organizations represented to those that coincidentally fit the definition of policy intermediaries. Further, the study assumed that participants were truthful in their responses and that their memories of their own experiences and impressions were accurate.

This research has certain delimitations. The study was conducted in only two states, and only a small number of local and state-level policy actors were interviewed within each state, a total of 35 in both phases of the study. Only six school districts and their teacher evaluation efforts were included in the study's first phase, three in each state. In addition, only small school districts were chosen as sites for the research. The

participating districts were not randomly selected, nor were they selected with the intent of representing geographic regions within the two states. Finally, individuals from only four state-level organizations in each state participated in the study's second phase. Despite these limitations and delimitations it is hoped that this multi-case study will help to generate a firmer theoretical foundation for examining how local actors interpret a new "wrinkle" in the accountability movement, and how those interpretations are affected by state and federal policy conversations.

Researcher role. It is important to specifically acknowledge my relationship to the topic of state teacher evaluation policy. From 2011 through 2014, I was employed as a specialist in the department of Policy, Research and Outreach at Education Minnesota, the state teachers union in my state. Much of my work there was devoted to providing information, resources, and support regarding Minnesota's teacher evaluation policy to educators around the state. In addition, my colleagues and I collaborated with personnel from other state-level organizations to develop shared, consistent guidance about Minnesota's policy. Thus, I can be considered a state policy actor in relation to teacher evaluation in Minnesota, and I have a kinship with the state-level policy actors who participated in this study's second phase. In addition, it is natural that I identify with educators at the local level because I spent the first 17 years of my career working in the public schools, first as a teacher and then as a district staff development coordinator. Because of my relationship to the topic and to the work of both local and state-level policy actors, it was especially critical for me to approach the research with honesty and integrity. Furthermore, because the study relied on interviews as the main source of data,

maintaining professional distance was important. As Brinkmann and Kvale state, "In interviewing, the importance of the researcher's integrity is magnified because the interviewer himself or herself is the main instrument for obtaining knowledge" (2015, p. 97).

The decision to conduct this research outside of my own state was an attempt to create the appropriate distance from at least the state context; it ensured that the political culture, policy histories, and major education stakeholders I would endeavor to research would be unfamiliar to me. In addition, choosing to study small school districts, most of which were rural, took me outside my own experience; only one of my 17 years in public education was spent in a small rural school district. During the other 16 years of my work in the public schools, I served in a medium-sized district of approximately 8,000 students within 30 miles of Minneapolis and St. Paul, Minnesota. Finally, my higher education experience (in teacher preparation) was at the University of Minnesota, an urban research institution, where most of my interactions with preK-12 personnel involved employees from urban and suburban school districts.

Definitions

To aid the reader in understanding the key terminology that is presented in this dissertation, the following definitions of terms are offered, as they are understood in this paper:

- Teacher evaluation: The process by which a teacher's practice is reviewed for the purpose of professional growth, employment decision-making, and/or educational accountability.

- Professional growth: Continuous improvement of practice.
- Employment decision-making: Personnel actions such as retaining, disciplining, or removing an employee; granting tenure; or rewarding strong performance.
- Educational accountability: The accountability of schools, districts, or states to external stakeholders including the public and agencies of government.
- Teacher effectiveness: A broad term recognizing the breadth of teaching responsibilities, this term refers to a teacher's ability to make contributions to positive academic, attitudinal, and social student outcomes, and to collaborate with others to promote student success and well-being.
- Teacher quality: A determination of teacher competence based on measurable teacher attributes. As framed by the literature review (see Chapter Two), the use of this term is related to No Child Left Behind (NCLB) expectations and other accountability demands.
- Policy actor: An individual or group that is actively involved in the policy process (Fowler, 2009).
- Policy interpretation: A distinct phase in the policy process, occurring after formal policy adoption and before full implementation. The conception of policy interpretation as a separate phase recognizes policy actors as learners and adapters of policy, not simply as individuals who execute policy mandates.
- Two main purposes of evaluation:
    - Formative evaluation: Focused on professional growth with an emphasis on development, improvement, and helping.

  - Summative evaluation: Focused on employment decision-making and educational accountability with an emphasis on administrative and personnel processes and meeting obligations to external stakeholders.

- Supervision: The responsibilities of school personnel who oversee both the teaching function *and* teachers as employees. Generally such responsibilities are assigned to the school principal.
- Role conflict: The understanding that an individual's behavior is based on his or her social role (Biddle, 1986).
- Political culture: The enduring political attitudes and behaviors associated with groups that live in a defined geographical context (Elazar, 1970).
- Policy intermediaries: Organizations or individuals who operate between policymaking and implementation, providing assistance to both policymakers and implementers by interpreting, negotiating, translating, adapting, and framing the policy and its requirements.
- Policy conflict: Conflict between a policy's goals, or between its goals and the reasons for its adoption.
- Density of policy collaboration: The frequency and degree of participation in policy activities, the intensity of cooperation across groups, and the level at which local policy actors interact with state-level policy actors in relation to the policy.

Overview of Chapters

This dissertation is comprised of six chapters, with the current chapter introducing the study. Chapter Two provides a review of the literature, addressing the literature on the

history of teacher evaluation practice and policy, teacher effectiveness, formative and summative evaluation, and state education policy. In Chapter Three, the study's methodology is described, with the various sections of the chapter focusing on research design and methods, sampling and data collection, and data management and analysis. Chapter Four presents the results of the Missouri case, while Chapter Five presents the results of the Oregon case. Each of these two chapters begins with background on the state, including sections on the education policy context, and teacher evaluation history. Next, each of the results chapters addresses the three school district cases. Finally, the state case is presented in each of these chapters. Chapter Six presents the study's findings and offers a discussion, including implications for policymakers and practitioners and recommendations for future research.

Chapter Two

Review of Literature

This review of the literature begins by documenting the history of teacher evaluation, from the beginnings of public schooling in the U.S. to the present day, exploring both teacher evaluation practices and policies. Next, the literature on teacher effectiveness is summarized in order to demonstrate how research has impacted teacher evaluation over time. The chapter then addresses a central concern articulated in both research and policy: the tension between the formative and summative purposes of teacher evaluation. In order to situate the problem of teacher evaluation within a policy frame, the relevant literature on state education policy is reviewed next. Finally, a summary of the chapter is provided and the study's conceptual framework is articulated. The literature reviewed here includes studies published in peer-reviewed journals, articles from practitioner-based publications, books and edited manuals, policy briefs, and policy guidance documents, all of which help to illustrate the state of teacher evaluation in the U.S. today.

History of Teacher Evaluation

Teacher evaluation in the U.S. has evolved from informal school inspections by members of the local community in Colonial times to the current laser-sharp policy focus on teacher evaluation policy at the both the state and federal levels. This section reviews the literature on the history of teacher evaluation practice and policy.

History of teacher evaluation practice: From community monitoring to collegial conversations. The literature documenting the history of teacher evaluation

practice is drawn from historical accounts of local practices, research on supervision, and documentation of best practice recommendations from journal articles and textbooks. The literature documenting the history of supervision is notably thin, making it difficult to document historical practices related to the evaluation of teaching. Indeed, Glanz (1995) articulated a concern about the void in historical studies on school supervision, stating, "supervision as a field of study and practice has escaped serious and ongoing investigation by educational historians (p.96)." Further complicating matters is the fact that teacher evaluation was under the purview of local authorities until the late twentieth century. Therefore, it must be understood that the trends articulated in the literature apply broadly to practices in U.S. schools. Local control of teacher evaluation practice resulted in wide variation between urban and rural areas and various regions of the country and undoubtedly reflected local values, personalities, and prerogatives.

The literature on teacher evaluation practices from the Colonial period through the nineteenth century is sparse but consistent in its depiction of the monitoring of teachers and schools. The Massachusetts School Law of 1647, an act that helped establish public schooling in the U.S., gave towns responsibility for the performance of schools. Community leaders held the responsibility for evaluating teachers, and school visits were the means by which the citizenry monitored the content of instruction, discipline, student progress, and maintenance of facilities – all responsibilities of the teacher (Eye, 1975; Tracy, 1995). Often, clergy were tapped to supervise schooling, since they were among the best educated. Teaching was not considered a profession, and since teachers were

thought of as servants of the community, it was in the community's interest to monitor schools and teachers (Marzano, Frontier & Livingston, 2011).

With the common school movement of the 1800s came the establishment of public school systems. County superintendents were often charged with visiting and reporting on schooling at the local level, and one focus of these visits was assisting teachers with their instruction. Thus, it is clear that efforts to address the formative purpose of evaluation – improving teaching – began early, though no commonly held criteria for pedagogical skill existed (Marzano, et al., 2011). In urban areas, school systems became more complex, and there was a recognized need for personnel to take on specialized roles in the management of schools. The role of the principal emerged in the latter half of the century, beginning in urban settings and eventually spreading to rural areas as well (Tracy, 1995).

A source of some confusion in the literature on teacher evaluation practice is the fact that the terms *supervision* and *evaluation* are often conflated. Supervision is a term most prevalent in the literature on educational administration, and this field's conception of the term has evolved and expanded over time. It is important to acknowledge that the reason that supervision and evaluation are often used interchangeably is that the same person (the principal) is often charged with both. Researchers agree that there has been a lack of consensus on the definition of the term *supervision* in the literature (Bolin, 1987; Alfonso & Firth, 1990; Glanz, 1995). According to Bolin, historically, supervision was defined as "inspection for the purpose of improving teachers" (1987, p. 370). Over time, additional emphases were added, such as helping teachers develop pedagogical skills,

establishing the appropriate conditions for learning, and guiding the implementation of curriculum. In 1982, the yearbook of the Association of Supervision and Curriculum Development (ASCD), the leading professional organization for educators engaged in supervision, stated, "supervision is a field broadly conceived, as a general school activity that encompasses a number of school roles and that includes virtually all of the activities of administrators and supervisors involved in the improvement of instruction" (Sergiovanni, 1982, p. vi). As stated in Chapter 1, for the purposes of this paper, *supervision* is defined broadly to mean the responsibilities of school personnel who oversee both the teaching function and teachers as employees, while *teacher evaluation* refers specifically to the review of teacher performance, which is one of the responsibilities of the supervisor.

In the latter part of the 1800s and into the twentieth century, the formative purpose of evaluation was emphasized, as those responsible for supervising teachers often took on the role of teacher trainer and curriculum developer (Bolin & Panaritis, 1992). School superintendents spent considerable time assisting teachers, and many held the view that teachers needed help with instruction and discipline (Glanz, 1977). Yet summative decisions about teacher competence were often based on moral standards and basic competencies. "Good" teachers possessed high moral standards, had solid basic skills, and were considered positive role models (Ellett & Teddlie, 2003).

In 1916, in the first edition of his book *Public School Administration*, Edward Cubberly described how Frederick Winslow Taylor's principles of scientific management, so influential in industry, could be applied to school administration.

Cubberly, viewed by many as the father of public school administration, thought scientific management would bring efficiency to the schools and scientific accuracy to the work of the school administrator, so he and others encouraged the use of rating scales for evaluating teachers (Fine, 1997; Marzano et al., 2011). This shift toward scientific approaches to teacher evaluation continued through the 1920s.

During the 1930s and 1940s school administrators began to view the improvement of instruction as a cooperative enterprise between supervisors and teachers. Recognition of the social and psychological needs of teachers influenced school decision-making and teacher evaluation practices, and the relationship of supervisors to teachers became less authoritarian and more collegial (Fehr, 2001; Tracy, 1995). With the 1950s came an era of research on teaching methods and an effort to link student outcomes with particular teaching behaviors. Writing about the history of supervision, Eye explains that by 1950 supervision was viewed as a "service that helps the teachers do a better job" (1975, p. 15). Classroom observations and checklists identifying favored teaching behaviors became common components of local teacher evaluation approaches (Ellett &Teddlie, 2003).

In the 1960s and 1970s, the emergence of clinical supervision as the dominant method by which school principals evaluated teachers marked the convergence of the collegial approach to supervision initiated in the 1930s and the checklist systems developed in the 1950s. Morris Cogan and Robert Goldhammer are credited with developing the clinical supervision model, characterized by reflective dialogue between the teacher and supervisor aimed at continuous improvement of practice, and the clinical

supervision methods of pre- and post-observation conferences and direct observation of classroom instruction continue to be practiced in schools today (Tracy, 1995; Marzano, et al., 2011). Clinical supervision was widely adopted, with 90 percent of supervisors using the model by 1980 (Marzano, et al., 2011). Marzano, et al. explain that over time supervisors moved away from the rich dialogue about teaching that the clinical supervision model intended and began to use the model's supervision cycle as a structure for summative evaluation, an application for which it was not intended. Thus, where Cogan and Goldhammer advocated clinical supervision as supportive and developmental, its application became focused on the high stakes decision-making.

It was during the 1980s that Madeline Hunter's model of supervision was introduced, and ultimately, this model was embraced as a major approach to teacher evaluation across the U.S. (Marzano, et al., 2011; Fehr, 2001). The centerpiece of Hunter's model was a seven-step framework for an effective teaching lesson. In addition, Hunter offered a set of templates for evaluating instruction. Hunter's recommendations for observation and coaching bore strong similarities to the cognitive coaching model of the 1960s and 1970s, though Hunter did not endorse a lengthy pre-conference, instead advocating professional development for both supervisors and teachers ahead of the evaluation cycle to develop a shared vocabulary of practice (Fehr, 2001). Praised for its focus on learning theory and effective teaching research as well as its influence on increasing professional conversations in schools, Hunter's model was criticized by some for being mechanistic and advocating a single acceptable approach to teaching (Gibboney, 1987; Fehr, 2001; Garman & Hazi, 1988).

Charlotte Danielson published her *Framework for Teaching* in 1996, and it has become a centerpiece of teacher evaluation practices in many schools and districts today, though this was not Danielson's original intent. Danielson set out to accomplish three goals in designing her framework, intended for use by pre-service teachers: to illustrate the complexity of teaching, to provide a language for collegial conversation, and to deliver a structure for teacher self-assessment (Marzano, et al., 2011). By 2011 Danielson's framework had been adopted or approved by nine states for teacher evaluation, and was being used in hundreds of school districts across the country (Elliot, 2011). In addition, countless states and districts use local adaptations of Danielson's work. Where Goldhammer and Cogan's cognitive coaching model articulated the steps in the supervisory process, and Madeline Hunter's model described the steps in the teaching process, Danielson's framework "sought to capture – in its full complexity – the dynamic process of classroom teaching" (Marzano, et al., 2011, p. 23). The Danielson framework addresses more than classroom practice; its four domains include planning and preparation, the classroom environment, instruction, and professional responsibilities (Danielson, 1996). Studies of the framework have shown that teachers find it to be credible and helpful to their teaching, and teacher scores on the framework are related to student achievement outcomes (Little, Goe & Bell, 2009).

History of teacher evaluation policy: From local responsibility to policies in every state. Teacher evaluation was the responsibility of local authorities from the beginnings of schooling in the U.S. until the latter part of the twentieth century. Neither the federal government nor states established policies that specifically addressed the

evaluation of teaching personnel. This began to change in the 1970s and 1980s as efforts to reform schools placed an emphasis on teaching quality. Prior to 1970, only four states had policies that required evaluating or assisting in-service teachers (Tracy & Smeaton, 1993).

The landmark report, *A Nation at Risk: The Imperative for Educational Reform*, issued by President Ronald Reagan's National Commission on Excellence in Education in 1983, ushered in a wave of efforts to reform public schools and colleges of education. Among its recommendations on teaching, *A Nation at Risk* called for improvements in teacher evaluation; it was one of the first widely publicized documents to advocate a model of evaluation associated with employment decisions for experienced teachers. The report stated: "Salary, promotion, tenure, and retention decisions should be tied to an effective evaluation system that includes peer review so that superior teachers can be rewarded, average ones encouraged, and poor ones either improved or terminated" (Gardner, Larsen, Baker & Campbell, 1983, p. 30). In that same year, a report from the Task Force on Education for Economic Growth, Education Commission of the States, echoed similar concerns and recommended improvements in teacher evaluation. The report, *Action for Excellence: A Comprehensive Plan to Improve Our Nation's Schools*, stated, "Ineffective teachers – those who fall short repeatedly in fair and objective evaluations – should, in due course and with due process, be dismissed" (Wise, Darling-Hammond, McLaughlin & Bernstein, 1985, p. 62). A "second wave" of reform began in 1986 with the release of a report by the Carnegie Forum on Education and the Economy (Furtwengler, 1995, p 2). This report and others of the time went further than *A Nation at*

Risk by pressing for career ladders, school restructuring and the professionalization of teaching.

The criticisms of teacher evaluation policies and practices had begun before *A Nation at Risk* and similar reports were issued, with the literature documenting a variety of concerns about the practice of teacher evaluation beginning in the 1970s. Existing systems were criticized as subjective, based on vague or non-existent criteria, and lacking in guidance for improvement (Ingils, 1970; Jones, 1972; Newton, 1980; Feldvebel, 1980; Kult, 1978). In addition, there were questions about the validity and reliability of teacher evaluation procedures and criticism about the inability of teacher evaluation systems to distinguish and reward various levels of performance (Feldvebel, 1980; Newton, 1980). In a study of teacher evaluation practices in 70 districts across 38 states, Ingils (1970) found a lack of objective criteria for the evaluation of teachers. Ingils' analysis also pointed to a lack of correlation between what was assessed on evaluation forms and the educational objectives of the districts and schools. In the *Handbook of Teacher Evaluation*, published by the National Council on Measurement in Education in 1981, Michael Scriven stated, "Teacher evaluation is a disaster. The practices are shoddy and the principles are unclear" (1981, p. 244).

The literature documented issues related to the tension between teacher evaluation's formative and summative goals and concerns about the practices of evaluators. Guthrie and Willower (1973) found that principal evaluations of teachers tended to be ritualistic rather than goal oriented, and that their comments on evaluation forms were positive and congratulatory, not constructive. The authors concluded,

"classroom observation is essentially impotent as a method of improving instruction" (p. 289). Felvebel asserted that a key problem with teacher evaluation was the expectation of evaluators to address both formative and summative goals, which call for "contradictory forms of social relationships between the evaluator and the teacher" (1980, p. 415). In a report on the status of teacher evaluation issued in 1982, Peterson and Kauchak argued that a major obstacle to effective teacher evaluation practices was the predominance of administrative (summative) evaluation, which cut out other purposes. These studies and others helped to create momentum for policy development around teacher evaluation.

The RAND study (Wise et al., 1985) investigated teacher evaluation practices across the U.S. and had a strong influence on policy development in succeeding years (Marzano, et al., 2011). The RAND research team identified four major problems with teacher evaluation: apathy or resistance on the part of teachers, a lack of consistency and uniformity in evaluation systems, a lack of training for evaluators, and a lack of competence and resolve on the part of principals to evaluate effectively. Explaining that principals valued collegial relationships with teachers and wanted to be viewed as "good guys," the researchers stated: "the conflict between the principal as instructional leader and evaluator has not been settled" (Wise, et al., 1985, p. 75). Further, the research team noted that teacher resistance to evaluation was partly due to their belief that evaluation results relied too heavily on the subjective judgment of the principal.

Partly in response to these criticisms, state teacher evaluation policies began to appear in the 1970s. Two of the earliest policies, passed in California and South Carolina, were fairly robust, specifying particular evaluation system components. In 1971

California Governor Ronald Reagan signed the Stull Act, which required evaluation of both probationary and tenured teachers, involvement of teachers organizations in the design of local systems, establishment of evaluation standards, and the inclusion of information on pupil progress in a teacher's evaluation (Popham, 1972). As it turned out, California's requirement that pupil progress be included in teacher evaluation system design went largely ignored for decades (McGlone, 2013). In 1979, South Carolina passed the Educator Improvement Act, which required staff development and observations of both probationary and tenured teachers (McDaniel, 1981). The realities of teacher evaluation practices were surprising, if not shocking, to some policymakers. In the early 1980s, Tennessee began its work on a new teacher evaluation system by investigating existing practices in districts across the state. State officials found that in most school districts, principals completed checklists on teacher performance at the end of the school year, "without ever having observed the teachers" (Furtwengler, 1991).

The knowledge base on teacher evaluation grew during the 1980s as reform efforts focused on teaching (Ellett & Teddlie, 2003), and increased policymaking related to teacher evaluation followed. Some of the early state policies were directed toward beginning teachers. Starting in 1980, beginning teachers in Georgia experienced the first statewide, competency-based evaluation system that tied performance to continuing certification (Ellett, et al, 1980; Larsen, 2005). In the mid-1980s, Connecticut established policies designed to support beginning teachers through observations of teaching, intensive mentoring supports, and professional development (Wilson, et al., 2001).

State policy development on teacher evaluation in the latter part of the twentieth century tended to follow national initiatives (Hazi & Rucinski, 2009). Indeed, there was a marked upswing in new state teacher evaluation policies in the 1980s in response to *A Nation at Risk* and the other reports discussed earlier. By 1985, 26 states had passed teacher evaluation requirements, with the vast majority having been enacted since the early 1970s, but most states left the design of evaluation systems to local authorities (Stiggins & Bridgeford, 1985). States continued to evolve their policies on teacher evaluation during the 1980s, with 44 states having some kind of teacher evaluation policy in place by 1990. Two studies of state teacher evaluation policies illustrate the degree of specificity in state mandates. In a 1993 study documenting teacher evaluation requirements across the U.S., Tracy and Smeaton reported that state teacher evaluation policies exhibited a range of levels of control from the state, with some policies providing general requirements for teacher evaluation and allowing local control over specifics, and others prescribing evaluation criteria and instrumentation. At the time, more states provided general rather than specific requirements (Tracy & Smeaton, 1993). Sixteen years later, in a fifty-state analysis of teacher evaluation policies, Hazi and Rucinski (2009) found that 58% of states exerted control over state teacher evaluation policy, from approving local plans, to specifying evaluation criteria, to requiring a specific evaluation procedures or instruments.

In 2002, President George W. Bush signed the reauthorization of the Elementary and Secondary Education Act, which became known as No Child Left Behind (NCLB). Among its key areas of focus, NCLB placed emphasis on teaching quality, ushering in

yet another wave of efforts to improve teaching in the U.S. In response to NCLB, the National Governors Association (NGA) issued a set of policy recommendations regarding teacher evaluation. In the report, *Improving Teacher Evaluation to Improve Teaching Quality*, NGA identified the following strategies for policy development: define teaching quality, focus on improving teaching, incorporate student learning, create accountability, train evaluators, and broaden participation in evaluation design (Goldrick, 2002). Now, another federal initiative had sparked discussion, and a national report helped direct the resulting policy initiatives. Hazi and Rucinski (2009) found that the majority of states adopted the NGA strategies. In addition, state oversight over teacher evaluation increased after NCLB.

A new wave of criticisms of teacher evaluation practices was issued in the 2000s. In addition to echoing many of the concerns articulated in the 1970s and 1980s, critics charged that teacher evaluations were superficial and capricious, failed to provide constructive feedback or address poor performance, and were rarely linked to student learning. Those who found fault with existing teacher evaluation practices blamed a host of factors, among them ambivalence from teachers unions, lack of training for evaluators, the indifference of schools to variations in teacher effectiveness, and the fact that NCLB based its definition of "highly qualified" teachers on credentials not performance (Weisberg, Sexton, Mulhern & Keeling, 2009; Marshall, 2005; Marzano, et al., 2001; Toch & Rothman; 2008). These concerns were addressed at the federal level during President Barack Obama's first term when two major policies – Race To The Top (2009) and the Elementary and Secondary Education Act (ESEA) Flexibility Program (2011) –

induced states to legislate teacher evaluation systems. The teacher evaluation requirements in these two federal policies were nearly identical (U.S. Department of Education, 2015b).

Currently, all 50 U.S. states and the District of Columbia have policies requiring teacher evaluation on the books, and all but a handful have newly adopted policies that were implemented by the 2014-2015 school year, (Center on Great Teachers & Leaders, 2013; Hull, 2013; Doherty & Jacobs, 2013; National Council on Teacher Quality, 2014b; Doore, B., Jorgensen, Doore, S. & Mason, 2013). In 2013, the American Institutes for Research reported that forty-nine states and the District of Columbia had updated their teacher evaluation systems in the past four years (American Institutes for Research, 2014). There is no doubt that the efforts of the Obama administration have encouraged a wave of reforms in state teacher evaluation policies. Further, the requirement in both RTTT and the ESEA Flexibility Waivers that a significant portion of a teacher's evaluation be based on student growth data has resulted in an upsurge in state policies that include measures of student growth or achievement; 40 states and the District of Columbia now have policies that include student achievement or growth as a teacher evaluation criterion (National Council on Teacher Quality, 2014a). Some researchers point out that such policies have been influenced by advancements in the ability to measure student growth and link the data to teachers (R. Campbell, Kyriakides, Muijs & Robinson, 2003; Hinchley, 2010).

How Research Has Affected Teacher Evaluation Practice and Policy

As Ellet and Teddlie state, "Historically, teacher effectiveness research has greatly influenced (and will continue to influence) teacher evaluation practices in the USA" (2003, p. 121). Thus, in order to understand the context of teacher evaluation policies and practices, it is necessary to review the research on teacher effectiveness. On the whole, this research has been conducted from a theoretical framework focused on determining teacher characteristics, qualities, and behaviors that influence positive student outcomes; in this somewhat problematic framework, the teacher is seen as the primary, if not only, influence on student learning. The connection between the teacher effectiveness literature and teacher evaluation policy is complicated by the continued lack of a clear, agreed-upon definition of effective teaching. Further, changing approaches to measuring teacher effectiveness have driven a narrowing of the uncertain definition that does exist. This narrowed definition is embodied in current policy discussions around teacher evaluation and can be seen as the new "public" definition of teacher effectiveness.

The teacher effectiveness literature: No clear definition of effectiveness. R. Campbell et al., (2003) trace the history of teacher effectiveness research to the early part of the twentieth century when researchers sought to establish a set of rational principles for effective teaching, mainly to help determine the content of teacher preparation programs. Such research was underway in the U.S., England, and other parts of Europe, and the definition that was identified – strong content knowledge, effective

lesson planning, understanding teacher-student interactions, and core teaching skills such as questioning and explaining – still rings true today (R. Campbell et al., 2003).

Beginning in the 1950s, influenced by scientific management and behaviorism, educational researchers began to look for links between teacher behaviors and student outcomes in an attempt to determine the most effective teacher behaviors, or methods of teaching. This research on effective teaching behaviors was the dawning of the process-product research phase, which was the predominant approach to studying teacher effectiveness from the 1960s through the 1980s. In simple terms, the process-product research tradition sought to determine a relationship between teacher behavior (process) and student learning (product) (Doyle, 1977; Gage & Needels, 1989; Westbury, 1988). The process-product research tradition typically involved observations of classroom activities focused on teacher and/or student behaviors as well as teacher-student interactions. Correlations between classroom observation data and student outcomes were sought in order to determine the most effective teaching practices. During the height of process-product research, behaviors of effective teachers that were gleaned from the research formed the basis of checklist systems, a common means of evaluating teachers (Ellett and Teddlie, 2003). These checklist systems called for observers to review a list of favored teaching behaviors, checking those that were observed during an episode of instruction. Though advocates of this approach praised its objectivity, it was criticized for focusing on basic teaching skills (Darling-Hammond & Berry, 1986), being procedural in nature, (Hallinger, Heck & Murphy, 2014), requiring inference on the part of the

observer, and using forms that did not undergo reliability testing (Peterson & Kauchak, 1982).

Process-product research faced substantial criticism. One critique had to do with the assumption of unidirectionality of causal influence from the teacher to the student; that is, the notion that teacher behaviors affect student learning and growth with no consideration of the effects of student actions on teacher behavior or other influences on student learning (Westbury, 1988; R. Campbell et al., 2003). Other critiques included concerns about: the reliability of observational methods; the reliability of product measures such as standardized tests; the limitation of investigating discrete teacher behaviors without addressing sequences of teacher action; and the generalizability of findings when results could be considered context- or content-specific (Gage & Needels, 1989; Westbury, 1988; Ornstein, 1995). In addition to these specific critiques, process-product research was included in broad critiques of research on teaching that came in the late 1980s. Research on teaching, which attempted to lay a scientific groundwork for the profession – going back to the early part of the twentieth century and peaking during the 1970s and 1980s – was considered by some to be inconclusive at best (Gage, 1989).

In the late 1980s, partly in response to criticisms of process-product research, an ethnographic model of research on teacher effectiveness was adopted, investigating the complexities of teaching, recognizing power structures, and exploring teacher thinking in addition to teacher action. Advocates of this model emphasized its utility for breaking down barriers between researchers and practitioners (Ornstein, 1995). Another research tradition that succeeded process-product research was the knowledge and beliefs

approach, which explored how student outcomes were influenced by teacher knowledge – of subject or pedagogy – and beliefs – such as high expectations for students. It is important to note that the interest of policymakers in research on school effectiveness rose during the 1980s and 1990s, causing a drop in the amount of teacher effectiveness research during that time frame (R. Campbell, Kyriakides, Muijs & Robinson, 2004).

About the same time, researchers began to explore teacher *effects* on student learning. In this approach, the teacher is a variable in an exploration of the effect of certain variables on student achievement (Ding & Sherman, 2006). A good teacher is identified as one who consistently produces gains in student learning, while a weak teacher is one who consistently fails to produce learning gains (Hanushek, Rivkin, Rothstein & Podgursky, 2004). William Sanders, who developed the Tennessee Value-Added Assessment System, was among the first to publish on this approach, and his work is often used to support the assertion that the classroom teacher is the single most important school-based variable in student achievement. Value-added modeling is a statistical advancement that makes it possible – generally utilizing standardized test scores – to isolate and measure a teacher's contribution to the growth in the achievement of a student (or group of students) over time. This contribution is considered the teacher's value-added. Value-added models (VAMs) are able to control for certain student characteristics, such as poverty or special education status. It should be acknowledged that there are multiple forms of VAMs, and methodologies differ (Baker, et al., 2010; Amrein-Beardsley, Collins, Polasky & Sloat, 2013). For purposes of this study, VAMs

are considered as the broad category of statistical measures of student growth that rely on standardized tests scores as the basis of the calculation.

Used in many states and school districts to meet policy mandates requiring the inclusion of measures of student growth in teacher evaluation systems, VAMs have received considerable criticism. These criticisms have included concerns about validity and uncertainty of results (Kupermintz, 2003) as well as lack of stability across statistical models, years, and the classes taught by an individual teacher (Baker et al., 2010). Although researchers disagree on the utility of VAMs as a component of teacher evaluation systems, there is consensus around the fact that "VAM measures are imperfect, and questions about bias are far from settled" (Corcoran & Goldhaber, 2013, p. 423). Indeed, in 2014 the American Statistical Association issued a statement advising caution in the use of VAMs in high-stakes decision-making, stating, "at best, most VAMs predict only performance on the test and not necessarily long-range learning outcomes. Other student outcomes are predicted only to the extent that they are correlated with test scores" (American Statistical Association, 2014, p. 4). Thus, the use of VAMs in teacher evaluation is highly controversial.

Challenges related to the implementation of VAM measures have been addressed, not only by researchers, but by the media as well. Educators and the public at large have been exposed to the debate about issues of fairness and reliability related to VAMs, along with concerns about measuring student growth for teachers in non-tested areas, misclassifying teachers, and introducing competition among teachers. Furthermore, VAMs are subject to the criticisms waged against standardized testing (since such test

scores are used in VAM calculations); among these criticisms are concerns over the narrowing of the curriculum, undue time devoted to test preparation, and increased challenges in hiring teachers for hard-to-staff schools (Jiang, Sporte & Luppescu, 2015; Scherrer, 2012; American Statistical Association, 2014).

Another body of literature on teacher effectiveness explores whether certain teacher characteristics – such as certification, education, and experience – influence student achievement. These studies are of interest because they explore factors that are typically addressed in teacher salary schedules. In most school districts, teachers are not employable without proper certification, and their placement on the salary schedule is determined by education (degrees and additional coursework) and experience. As pay-for-performance approaches have grown in popularity, these studies have become more and more salient in policy discussions. Some recently adopted state teacher evaluation policies eliminate traditional salary schedules, linking pay to evaluation results (Harrision & Cohen-Vogel, 2012).

In a meta-analysis of studies of the effects of particular teacher characteristics on student achievement, Wayne and Youngs (2003) found that students learn more from teachers with higher scores on tests, such as licensure exams and tests of verbal ability; that students learn more from teachers who attended highly rated colleges; and that high school students learn more in mathematics courses from teachers with more coursework or advanced degrees in mathematics. (Only in mathematics did additional coursework or advanced degrees have an effect on student learning, and only at the high school level.) The analysis did not determine a positive relationship between student achievement and

teaching experience. Podgursky and Springer (2007) summarized studies from the 1990s and 2000s that measured value-added teacher effects on student learning and their link to teacher characteristics. The authors state that estimated teacher effects are, in the main, unrelated to teacher characteristics such as certification, education, test scores, and teaching experience (beyond the early years). According to Staiger and Rockoff, studies of teacher effects on student learning show that "Teachers improve substantially in their first few years on the job" but not beyond (2010, p. 102).

The historic inability of research to demonstrate a solid link between teacher characteristics and student achievement has resulted in calls for results from robust teacher evaluation systems to replace characteristics in determining teacher salary. However, some recent studies have demonstrated that teacher experience *is* linked to student success, including evidence that experienced teachers bring higher rates of student achievement in mathematics (Wiswall, 2013), that teachers continue to improve throughout their careers (Papay & Kraft, 2014), and that experienced teachers are more successful when measured both by student performance and by improvements in student behavior (Ladd & Sorenson, 2014).

R. Campbell et al. (2003) assert that throughout the history of teacher effectiveness research, the definition of effectiveness has been determined to a substantial degree by the available means of measuring it. The authors state: "It is difficult to avoid a sense of the horse and the cart being in the wrong places; the technology of measurement has been creating the concept of effectiveness rather than the concept requiring an appropriate technology" (p. 352). In addition, R. Campbell and colleagues express the

criticism that teacher effectiveness research has succeeded in distinguishing ineffective from average teaching, but not in distinguishing highly effective from average teaching. The authors' central criticism is the concern that teacher effectiveness research has identified a generic model of effectiveness, one that does not recognize variations in contextual factors such as school environment, student ability, community demographics, or even grade level and subject.

Recent views on teacher effectiveness and links to policy. Teacher effectiveness research and policy development in the U.S. is situated within an international environment also focused on teacher quality. In 2005, the Organization for Economic Co-operation and Development (OECD), an organization of 30 countries working together to address challenges of globalization, issued a report on effective teaching based on a review of research and policy across the globe. The report addressed a broad range of issues, including the attractiveness of teaching as a career, teacher education, recruitment and selection, induction, and retaining teachers in the field. Among its broad findings, the reports states that, "of those variables which are potentially open to policy influence, factors involving teachers and teaching are the most important influences on student learning" (McKenzie & Santiago, p. 26). The report goes on to state that raising teacher quality is essential for raising student achievement. Of course teacher evaluation is only one means of addressing teacher effectiveness through policy.

In a 2008 research synthesis issued by the National Comprehensive Center for Teacher Quality, Goe, Bell, and Little state that there continues to be a lack of agreement on a definition of effective teaching and how to measure it. This problem has been

identified repeatedly in the literature (Cruickshank & Haefele, 1990; Ding & Sherman, 2006; Looney, 2011). Goe and colleagues distinguish three common approaches to measuring teacher effectiveness: measurement of inputs, a teacher's characteristics and credentials; measurement of processes, classroom interactions between the teacher and his or her students; and measurement of outputs, data on student achievement or other educational goals. Goe et al. note that defining effectiveness through examining credentials has come to be termed "teacher quality" in policy discussions, due to the No Child Left Behind (NCLB) requirement that all teachers meet the law's definition of "highly-qualified," which emphasizes credentials. Goe, et al. point out that the definition of teacher effectiveness has narrowed in recent years: "Increasingly, policy conversations frame *teacher effectiveness* as a teacher's ability to produce higher than expected gains in students' standardized test scores" (p. 5). The authors argue that this narrowing of our understanding of teacher effectiveness takes the problematic assumption of causality too far, ignores important ways in which teachers contribute to student success, and fails to address influences on student achievement that go beyond the teacher.

In a 2010 policy brief issued by the National Education Policy Center, Hinchley asserts that the emphasis on test scores in measuring teacher effectiveness is a result of factors having more to do with convenience than anything else. The author posits that the current focus on student achievement scores for measuring effectiveness is due to the availability of more student achievement data due to NCLB requirements, the recent emphasis by the Obama administration on using such data to evaluate teachers, and the perception of non-statisticians that statistical analyses are objective and reliable.

Similarly, Yeh (2009) offers a multidimensional definition of teacher *quality*, but he points out that research on this construct is limited to a narrow sub-set of aspects of teacher quality, those for which proxy data are available. These data include student test scores, scores on teacher licensure exams, and teacher characteristics such as experience and level of education.

Many researchers express concerns about defining teacher effectiveness based on student achievement scores alone. For example, R. Campbell et al. (2003) point out that teachers spend less than half their time in classroom instruction and that, during school reform efforts, the time outside the classroom is especially important. Fenstermacher and Richardson (2005) argue that effective teaching requires both quality methodology and student success. Further, they lament the U.S. tendency to "think of learning in terms of a Lockean *tabula rasa*, a blank slate wherein the teacher simply writes the content to be learned on the blank slate of the mind contained within a passive, receptive student" (p. 190). Instead, the authors argue, many factors combine to produce student success, including the student's own motivation and the school environment for learning.

Teacher effectiveness defined: The consensus that does exist. Recognizing the difficulties associated with research on teacher effectiveness and the acknowledged flaws that have been introduced into the discussion by measurement issues and policy frameworks, it is nonetheless useful to summarize the literature's depiction of effective teaching practice. Despite the concerns discussed to this point, there is in fact some consensus on what effective teaching looks like.

Based on their synthesis of the literature spanning the late 1970s through the 2008, Goe and her colleagues (2008) offer a comprehensive definition of teacher effectiveness that they assert can be useful in designing teacher evaluation systems. This definition recognizes that the teacher's role goes far beyond delivering instruction, and it attempts to broaden the narrowed definition of effectiveness that recent measurement advances and policy conversations have produced. The authors' definition of effective teaching includes the following concepts: holding high expectations for all students; making contributions to positive academic, attitudinal, and social outcomes for students; utilizing effective planning, instructional, and assessment techniques and adapting instruction as needed; encouraging a school and classroom climate that honors diversity and civic-mindedness; and collaborating with educators, families, and education stakeholders to promote student success and well-being. This definition is consistent with other definitions of effective teaching drawn from reviews of the literature (Stronge, 2007; Looney, 2011).

Teacher Evaluation's Two Purposes: Formative and Summative

Researchers, policymakers, and practitioners have identified a range of goals for teacher evaluation, and there is consensus on two central purposes: the formative purpose and the summative purpose. For decades the literature on teacher evaluation has reflected disagreement about whether these two purposes can be addressed within the same evaluation system. This section will outline the controversy and describe how it has been addressed in teacher evaluation policies and practices.

The Handbook of Teacher Evaluation published by the National Council on Measurement in Education in 1981 provides the following useful definition of formative and summative teacher evaluation:

> Formative teacher evaluation helps teachers improve their performance by providing data, judgments, and suggestions that have implications for what to teach and how. On the other hand, summative teacher evaluation serves administrative decision-making with respect to hiring and firing, promotion and tenure, assignments and salary. (Millman, 1981, p. 13)

Here we see that the emphasis of formative evaluation is on professional growth, helping teachers improve their practice, while the emphasis of summative evaluation is on employment decision-making. This definition of summative evaluation is limited because it ignores a major goal of summative evaluation, one that has become more prevalent over time. According to Darling-Hammond, Wise, and Pease (1983), an important aspect of the summative purpose of evaluation is accountability related to school improvement and status, what might be termed educational accountability. Both employment decision-making and educational accountability are included in this study's conception of summative evaluation. In the literature on teacher evaluation, formative evaluation is referred to as the developmental, improvement, helping, or growth-oriented purpose, while summative evaluation is frequently referenced as the accountability, administrative, personnel, or accountability-based purpose (Duke, 1985; Colby, Bradshaw & Joyner, 2002; Hazi & Rucinski, 2005).

The tension between the formative and summative purposes of personnel evaluation is not unique to education. The human resources literature has documented concerns about conflict between the "evaluative" and "coaching" purposes of performance appraisal for decades, with some researchers advising a separation of the two functions (Campbell, D. J. & Lee, 1988, p. 308). As far back as 1965, Meyer, Kay, and French recommended an approach to performance appraisal in industry where periodic coaching sessions focused on goal setting and problem solving with no summary judgments made, and with salary discussions held separately. This approach was designed to reduce employee defensiveness and motivate performance improvement.

The inherent conflict between the two main purposes of teacher evaluation is confounded by contradictions in how teacher evaluation policies are framed. While the majority of U.S. states indicate that professional growth for teachers – the formative purpose of evaluation – is central to their teacher evaluation policies (Minnici, 2014), the political rhetoric accompanying the development and enactment of these policies is often centered on the summative purpose. As described in Chapter One, some scholars view new teacher evaluation policies as part of a teacher accountability movement that embraces market-based reforms, promotes the use of test scores in decision-making, and places the teacher at the center of discussions about problems in public education (Ravitch, 2012; Connell, 2013; Shiller, 2015; Kumashiro, 2012). These scholars assert that in the narrative of the teacher accountability movement, a "'bad teacher' discourse" depicts teacher evaluation systems as a way to remove low-performing teachers (Stern, 2013, p. 213). This narrative promoting the summative purpose frames low student

achievement outcomes as a direct result of teachers' actions and blames underperforming or negligent teachers for impeding student learning (Harrison & Cohen-Vogel, 2012).

Concerns about incompatibility. In a 1988 article titled "The Dysfunctional Marriage of Formative and Summative Teacher Evaluation," W. James Popham articulates the challenge of blending the two purposes of teacher evaluation, asserting that they must be kept separate. His views represent those of many who find the formative and summative purposes of teacher evaluation incompatible. Popham states that the attempt to address both the formative and summative purposes of teacher evaluation in a single system is a major shortcoming in teacher evaluation design. He explains that the two approaches cannot be merged because their aims are so different; where formative evaluation aims to improve a teacher's performance, summative evaluation aims to identify weak teachers and dismiss them. He states, "In essence, the distinction is between 'fixing' and 'firing' the teacher" (p. 270). According to Popham, when school principals attempt to perform evaluations that serve a dual purpose, teachers are likely unwilling to reveal weaknesses or even acknowledge them. He also notes that principals tend to rate teachers highly because of their desire to be supportive, confounding the summative purpose. Popham recommends that different individuals carry out the two evaluation functions, and that the records be kept separate.

As teacher evaluation became a prominent policy focus in the 1980s, other researchers documented the challenge of the dual purposes of evaluation. Darling-Hammond et al. (1983) asserted that the conflicting purposes call for different types of data, stating that data for improvement must be rich and descriptive in order to prompt

growth and change, while data for accountability must be objective, standardized, and externally defensible. Stiggins and Bridgeford (1985) voiced the concern that due to the conflicting nature of most teacher evaluation systems' dual purposes, the promise of formative evaluation often goes unfulfilled. In a case study of four local evaluation systems, Stiggins found that evaluators lack skills in communicating with teachers about evaluation results; that there is a lack of time for follow-up, limiting feedback to teachers; that goals and processes supporting formative evaluation are unclear, making the promotion of professional development difficult; and that trust is lacking, partly due to the lack of involvement of teachers in system design, resulting in skepticism on the part of teachers toward the entire evaluation exercise.

In a 1995 article documenting the history of the field of supervision, Glanz termed the dilemma faced by principals "the inservice education and evaluation conflict" (p. 101). The author goes on to describe the conflict in principals' divergent responsibilities, explaining that in order to lead effectively, principals must develop supportive, collegial relationships with teachers for the purpose of assisting in instruction. However, professional relationships between principals and teachers are jeopardized when the principal must fulfill evaluation duties. Glanz asserts that this dilemma has not received enough attention in the literature and is therefore unresolved. The RAND study, mentioned earlier, concluded that principals were challenged by the competing responsibilities of instructional leader and evaluator, causing them to be ineffective in evaluating teachers (Wise et al., 1985).

Research published more recently continues to echo concerns about the conflicting purposes of teacher evaluation. In their 2002 review of the teacher evaluation literature, Colby et al. indicate that, though difficult, it is possible for teacher evaluation systems to effectively address both formative and summative purposes. However, Taylor and Tyler (2012) note that the summative function of performance appraisal systems can subvert the formative function because the link to personnel decisions may cause employees to view the evaluation process as judgmental and punitive. In response to the recommendations of Popham and others to split the formative and summative roles between different evaluators, Milanowski (2005) studied whether split roles led to better evaluation outcomes. He did not find that teachers were more open to discuss difficulties or accept performance feedback when roles were split. However, the study's findings suggest that it is important to ensure that developmental assistance is provided via the teacher evaluation system, whether the roles are split or not.

Citing studies demonstrating that teacher evaluation systems focused on professional learning contribute to both increased teacher effectiveness and student learning, Looney (2011) acknowledges the continued tension between the formative and summative purposes of evaluation and calls for an appropriate balance in evaluation system design. In a review of international research focused on the formative purpose of evaluation, Looney states that teachers benefit from timely, specific feedback in a supportive, collegial environment.

In 2013, twenty-five years after Popham's previously cited article was published, he reiterated his views on the importance of separating the formative and summative

functions of teacher evaluation. In his book *Evaluating America's Teachers: Mission Possible?*, Popham articulates four potential implementation mistakes "waiting to be made" as school districts and states respond to recent federal initiatives on teacher evaluation (p. 12). One of these potential mistakes is "confusing the roles of formative and summative teacher evaluation" (p. 17). Popham illustrates the necessity of balancing the two purposes if they are to be attempted in a single system:

> . . . as heart-warming and wonderful as improvement-focused formative teacher evaluation can appear, care must be taken so that its presence does not diminish the accuracy of summative teacher evaluation. Similarly, we dare not let the presence of summative teacher evaluation diminish the potency of formative teacher evaluation – a process from which most teachers and their students benefit substantially (p. 18).

Indeed, despite the concerns of Popham and others, more and more states and districts are adopting teacher evaluation policies that require the fulfillment of both broad purposes, most often conducted by a single evaluator, the building principal.

The promise of the formative purpose. Does teacher evaluation achieve the promise of its formative purpose, improving teaching practice? Researchers have had difficulty demonstrating a direct link between teacher evaluation and improved teaching (Colby et al., 2002). Duke (1995) summarizes the concerns of scholars who argue that evaluation intended to meet formative goals requires trust and risk-taking, realities that are difficult to achieve when the specter of accountability looms. Conley and Glasman (2008) stress that even in teacher evaluation systems that emphasize the formative

purpose, it is difficult to reconcile the organizational goals of improving school effectiveness with teachers' professional goals of improving individual teaching practice; indeed a central fear of teachers about evaluation is not receiving substantive feedback about one's teaching. Sinnema and Robinson (2007) argue that teacher evaluation has not achieved its purpose of improving teaching and learning because of issues such as the legal complexities and risks associated with teacher evaluation and the concern of both teachers and school administrators about the validity of evaluation criteria. In a series of linked empirical studies investigating teacher evaluation practices in 17 New Zealand elementary schools, Sinnema and Robinson find that there is little focus among educators on the teaching-learning relationship.

Some scholars do report positive findings related to the formative purpose of teacher evaluation systems, though studies exploring this question are few, and more research is needed. In the 1980s, Iwanicki and Rindone reported on the implementation of teacher evaluation policies in Connecticut over a twenty-year span (1985). Although they did not specifically investigate the question of whether teacher evaluation resulted in improved teaching and learning, the authors expressed the belief that Connecticut's teacher evaluation system had improved the process of teaching and learning in many schools. They estimated that teacher evaluation in Connecticut was "having an impact on the professional growth of teachers in 60-70 percent of the classrooms in 60-70 percent of the school districts" (p. 96). Iwanicki and Rindone expressed the concern that teacher evaluation cannot have a positive impact on teaching and learning in those schools where

administration and teachers do not believe that professional growth is a goal of the process.

More recent studies by Ovando (2001), Munson (1998), and Kleinhenz and Ingvarson (2004) demonstrate the promise of teacher evaluation for professional development, with each of the studies suggesting that observation and feedback from peers may help promote the formative purpose. Ovando (2001) investigated teacher perceptions about the effectiveness of a teacher evaluation system for enhancing professional growth and found that support from principals and others in supervisory roles is critical to the effectiveness of the system. System elements that teachers find beneficial for promoting professional growth include time to observe other teachers' classrooms and meet with peer coaches, collegial conversations about student learning, targeted professional development, feedback on performance, and self-evaluation. A study by Munson (1998), which also explored teachers' perceptions, concludes that peer review is more effective than administrative observation for enhancing teaching skills, with teachers reporting less anxiety and discomfort in a peer observation process. Researchers in Australia found that, although new teacher evaluation policies had not produced the promised improvements in teaching and learning broadly, teachers believe that the evaluation system promotes professional growth. This success was attributed to the direct link between the teacher evaluation system and the technical core of teaching, with evaluations performed by expert peers (Kleinhenz & Ingvarson, 2004).

Three recent studies underscore the promise of teacher evaluation to improve teaching and learning. In research comparing teacher performance as reflected in value-

added measures of student learning before, during, and after a year of teacher evaluation activities, Taylor and Tyler (2012) found that teacher evaluation improved teacher performance. This improvement was reflected in the performance of individual teachers during the evaluation year and was even stronger in succeeding years, improving performance "in a lasting manner" (p. 3647). The authors find that the teachers whose performance improved most after the evaluation year were those whose performance was weakest before experiencing the teacher evaluation system. Although they are unable to pin point the exact cause of the improvement, the authors attribute their findings to the teachers' experience learning about their performance and developing new skills.

In a 2014 article on the impact of human resource management and pay-for performance reforms on student achievement, Sojourner, Mykerezi, and West investigated Minnesota's Alternative Teacher Professional Pay System, also known as Quality Compensation, or Q-Comp. In schools that adopted Q-Comp, reforms included changes to professional development, teacher evaluation, and compensation systems. The study finds that the reforms are associated with an increase in student achievement. Although the authors are not able to "unbundle which pieces of the reform matter" (p. 25), they note that the findings are consistent with research showing that teacher evaluation can have a positive impact on student outcomes.

Jiang and colleagues (2015) investigated the perceptions of teachers about a new teacher evaluation policy in the Chicago Public Schools. Utilizing two years of survey data and interviews of a random sample of teachers, the researchers demonstrate that teachers are positive about the new teacher evaluation system overall, yet they express

concerns about the use of student growth measures in rating their performance. Specifically related to the question of evaluation's formative purpose, the study finds that teachers view the system to be fair and the feedback useful, providing guidance on how to improve instruction. The authors conclude that Chicago's teacher evaluation reform holds promise.

Potential role conflict for evaluators. As teacher evaluation policies are implemented, there is a potential that evaluators, usually building principals, will have difficulty managing the dueling purposes of evaluation, formative and summative. Role theory helps us understand this potential conflict. *Role theory* is the understanding that an individual's behavior is based on his or her social role. Based in sociology, role theory is applied to the broad range of social roles, from family roles (mother, child) to work roles (worker, manager).

It can be argued that in addressing the competing demands of formative and summative evaluation, principals may face *role conflict*. The literature on role theory explains that *role conflict* can result when an individual is faced with incompatible sets of expectations (Biddle 1986). As this literature review demonstrates, many scholars view the formative and summative purposes of teacher evaluation to be incompatible; this incompatibility applies both to the question of achieving formative and summative goals within the same system and to the question of the same evaluator being able to achieve both purposes. In addition to the articles throughout this literature review that have cited difficulties for supervisors related to the conflicting purposes of formative and summative evaluation, two studies specifically explore role theory and document role conflict among

principals related to teacher evaluation. A study by Gmelch and Torelli (1993) demonstrated that school administrators experience stress when there are incongruous expectations associated with the principal role, including conflict when the role of serving as a support to teachers is incompatible with the role of evaluating teacher performance. In a study on collaborative school leadership, Clift, Johnson, Holland, and Veal (1992) found that principal acceptance as a building leadership team member was affected by the teacher evaluation system, which was tied to salary and career advancement rewards. The authors state, "This sometimes led to role conflict when principals who were team members also served as evaluators of teacher performance" (p. 901). In addition, teachers in this study feared that there could be negative repercussions in the teacher evaluation process if they disagreed with the principal in team meetings.

Occupational role conflict is a particular type of role conflict that involves membership in multiple groups that hold conflicting expectations for the focal person. Stryker and Macke (1978) explain that, when individuals belong to a profession, the values of the profession may conflict with the expectations received by the focal person within the work role. Because the field of educational administration values the formative purpose of teacher evaluation, a principal receives the message from the profession that he or she is to act as a supportive guide when conducting teacher evaluations. However, in response to teacher evaluation policies that emphasize summative evaluation, the principal receives a message that he or she is to act as a high stakes evaluator when evaluating teachers. Thus, there is a potential that the principal will experience role conflict due to conflicting, incompatible sets of expectations.

State Education Policy

As suggested in the previous sections, the debates about teacher effectiveness and how to improve it have driven a great deal of research. While the connections between research and policy are typically fragile, this section is linked to the previous ones because the policy debates also reflect varying concerns with ensuring teacher effectiveness through high stakes evaluation and formative interventions. Thus, this section reviews the research on state education policy, with emphasis on the aspects most critical to understanding how teacher evaluation policy has played out over time and how research has reflected the public policy conversation.

History and context: Reform efforts and increased state capacity. Under the Tenth Amendment to the U.S. Constitution, domain over education is reserved to the states. Historically, states delegated much of the power over education decision-making to local authorities, which controlled schools both politically and fiscally. This began to change in the latter part of the twentieth century when states began to take back control of education policy (Fowler, 2009; Louis, 1998; Superfine, 2005). Today, the state is viewed as a legitimate actor in education policy possessing authority over local school districts, and this authority is rarely questioned (Louis, Thomas & Anderson, 2010).

Education policymaking in the states increased dramatically during the 1980s. Mazzoni (1994) proposes a political interpretation to explain the influence of the Reagan administration on education policy, stating that Reagan sought a devolution of educational authority and accomplished it through advancing symbolic policy and rhetoric, moving away from substantive policy at the federal level. 1983's *A Nation at*

Risk, with its imagery of the U.S. as a nation in peril due to its failing educational system, led to intensified media coverage of education and pressure on state policymakers to take action. States responded by enacting policies to strengthen the educational system. Mazzoni notes that issues of accountability had begun to be addressed in state education policies in the 1970s; thus, *A Nation at Risk* can be seen as having shepherded a movement that was already underway.

According to Louis (1998), three themes were woven into the educational reform efforts that began in the 1980s: ensuring equity by meeting the changing needs of students and families, making the educational system accountable to stakeholders, and safeguarding U.S. competitiveness by preparing students for a global economy. Louis points out that federal and state policy reforms, including the emphasis on standards and accountability, are at odds with the tradition of local control by elected school boards. Thus, the reform movement that began in the 1980s made the education policy environment more complex.

In addition to efforts toward reform, Rosenthal and Fuhrman (1981) articulate two accompanying reasons for the general increase in education policymaking that began in the 1980s: the changing nature of politics in education and significant changes in the structure of state legislatures. According to Rosenthal and Fuhrman, the role of states in education policy increased as states intervened in conflicts between various local education stakeholders – teachers, administrators, and school boards. As a result, policymaking efforts began to shift upward from the district to the state level. Also, the reality of an overburdened property tax system and court decisions on education finance

contributed to this shift. At the same time, the capacity of state legislatures increased, with standing committees gaining power and legislators gaining expertise in the issues addressed by the committees they served. According to Mazzoni (1994), this increase in the capacity of state governments to address complex policy issues had increased over time, so that by the 1980s when the calls for education reform sounded, states were ready to respond. The result was that the 1980s saw not only more state action on education, but also an increase in the proportion of education funding provided by states (Fuhrman & Elmore, 1990).

Increasing influence of federal policy on states. In 1994 another shift occurred when President Bill Clinton signed Goals 2000 into law. To many, this policy represented an unprecedented intrusion of the federal government into education policy. The policy attempted to promote reform on a national scale, incentivizing states around standards, assessments, and accountability measures (Superfine, 2005). Goals 2000 ultimately failed to maintain bi-partisan support and many of its harshest critics attacked the federal involvement in education policy with which it was associated. The program was plagued by political difficulties, poor enforcement mechanisms, and lack of state capacity for implementing the reforms in a manner coherent with other reform initiatives (Superfine, 2005). Yet Goals 2000 can be seen as having set the stage for expanded federal involvement in education in the ensuing decades.

In 2001, President George W. Bush signed the No Child Left Behind Act (NCLB) into law, reauthorizing the Elementary and Secondary Education Act (ESEA) first passed in 1965. Throughout its history, ESEA has been focused on ensuring equity in public

education. With NCLB, there was a shift from a focus on equality of opportunity to a "much more ambitious target of equalizing educational achievement" (Berry & Herrington, 2011, p. 273). The legislation required states to enact certain reforms in order to receive federal funding, among them establishing curriculum standards, developing student assessments, and reporting performance by subgroups of students as defined by their ethnicity, race, socioeconomic status, and other factors. Schools that failed to meet targets faced a variety of interventions, with the state expected to monitor schools' performance and enact the interventions. This linking of federal funds to state education policy has been highly controversial (Berry & Herrington, 2001). Regarding teachers, the policy required state and local authorities to ensure that teachers were "highly qualified." The definition of a highly qualified teacher was based mainly on credentials (Palardy & Rumberger, 2008).[3]

President Barack Obama's administration has sustained the expanded federal role in education through its policies, including the Race to the Top program (RTTT) and ESEA Flexibility Waivers. RTTT is part of the American Recovery and Reinvestment Act of 2009. A competitive grant program to state education agencies, RTTT continued the federal government's emphasis on standards and assessments and introduced a new emphasis on ensuring educator effectiveness. In addition to other requirements, states receiving RTTT funds were expected to develop systems for evaluating principals and teachers, with "effectiveness" based not on credentials but on evaluation results. In addition, as stated earlier, evaluation systems were to include data on student growth as a

[3] According to the federal definition, teachers are considered highly qualified when they hold a bachelor's degree, full certification or licensure, and prove they possess content knowledge in each subject they teach (U.S. Department of Education, 2015c).

"significant factor" in an educator's evaluation (U.S. Department of Education, 2015c). RTTT had the goal of moving from a system of sanctions under NCLB to a system of incentives to encourage innovation (McGuinn, 2012b). However, because of the prescriptive nature of the requirements, it can be viewed as a set of mandates, making it "more similar to NCLB than different" (Berry & Herrington, 2011, p. 288).

Using his authority to waive provisions of congressional legislation under the Department of Education's purview, Obama's Education Secretary Arne Duncan announced the ESEA Flexibility Waiver program in 2011 (Barron & Rakoff, 2013). Stating that NCLB was "far too punitive" and "far too prescriptive," Duncan emphasized the need to provide states more flexibility in setting high standards for improved teaching and learning (Office of the Press Secretary, 2011). A central motivation for the development of the waiver program was the failure of a divided congress to reauthorize ESEA. In addition to providing flexibility around some of the major provisions of NCLB, the waiver program created new requirements around college- and career-ready standards and differentiated accountability systems for schools, and – consistent with RTTT – evaluation of principals and teachers. Although the waivers relieve states from some of the most problematic aspects of NCLB, they are tied to the core purposes of the original act, a requirement of the statutory rules under which Duncan exercised his authority to issue the waivers (Barron & Rakoff, 2013, p. 336).

As stated, RTTT and the ESEA Flexibility Waivers sustained the increased involvement of the federal government in education policy that began with Goals 2000 in the mid-1990s. Where RTTT incentivized states with grant funds to adopt innovations

embraced by the Obama administration, the waivers provided the carrot of flexibility from certain provisions of NCLB in order to encourage state alignment with the administration's favored approaches to school reform. In terms of the number of states participating, RTTT's influence on state policy was limited compared to that of the ESEA Flexibility Program. To date, nineteen states have received RTTT funding (U.S. Department of Education, 2015c), while 43 states have received ESEA Flexibility Waivers[4] (U.S. Department of Education, 2015b).

Shifts in the locus of control: What does it mean? The growth in first the state and then the federal role in education can be seen as a shift away from local control of schools, but some researchers challenge that view. According to Fuhrman and Elmore (1990), the increased involvement of states that was spurred by the reform movement of the 1980s did not decrease the involvement of local authorities; instead, policymaking increased at both levels. Relations between various levels of government in policymaking have often been conceived as a zero-sum game, wherein as one level expands its policymaking activity, the other levels necessarily lose power and control. But during the 1980s, as state education policymaking increased, local activism in education increased as well, with districts anticipating state policy actions, enacting policies ahead of states, and using state policy initiatives to advance local objectives (Fuhrman & Elmore, 1990). Marsh and Wohlstetter (2013) characterize relations between federal, state and local governments as multidirectional, with influence going in multiple directions at once. The authors assert that, where the state or federal government may appear to have more

[4] At the time of this writing, the U.S. Congress is debating the reauthorization of ESEA. Both the House and Senate have passed bills, and neither would require states to evaluate teachers. If President Obama signs a bill reauthorizing ESEA, the ESEA Flexibility Waiver program would end (Camera, 2015).

control over the classroom as new policies are enacted, during the implementation phase local authorities reign, often shaping policy interpretations and influencing new policy directions.

Mehta (2013) argues that the shift upward in policy influence from the local, to the state, to the federal level that began in the 1980s is the result of a paradigm shift. Echoing Louis' (1998) articulation of the prominent themes of education reform, Mehta states that the new paradigm is characterized by four factors: a focus on standards and accountability that grew out of an increasing sense that the role of public education is to ensure economic stability; the calls for system-wide reform; the argument that schooling, not society, is responsible for improving student learning; and the belief that the best measure of effectiveness is test scores. Mehta credits *A Nation at Risk* for giving shape to this new paradigm that had been forming since the 1970s. According to Mehta, the new paradigm drew support from both political parties and challenged the power of teachers unions by casting them as defenders of the status quo. Mehta's study of three states, with various political characteristics and education policy histories, shows how states embraced standards-based reform and demonstrates how state-level policymaking activity ultimately influenced federal education policy. Mehta asserts that the influence of the new paradigm, "crystalized in *A Nation at Risk*" (p. 306), demonstrates that politics can be shaped by powerful paradigms.

The influence of political culture. Although federal influence is increasing, federal policies are filtered through specific legislative and administrative decisions that are developed at the state level. These decisions are influenced not only by the

parameters of federal laws (NCLB) and incentive programs (RTTT), but also by the deeper history of educational policymaking within each state, referred to as political culture.

Elazar (1970) defined political culture as enduring political attitudes and behaviors associated with groups that live in a defined geographical context. Elazar identified three political cultures that are associated with early state history: moralistic, traditionalistic, and individualistic. According to Elazar, in traditionalistic political cultures, dominant in the South, the established elite provides leadership, and the overriding political goal is maintaining the established order. In traditionalistic states, government is seen as a positive force, as long as it helps maintain the status quo. Moralistic cultures are dominant in the Northeast and in areas of the country settled by immigrants from Scandinavian countries, such as Minnesota and Wisconsin. In moralistic states, there is a belief in politics as an activity centered on the public good, and there is widespread public participation in voting and civic affairs. Individualistic political cultures view government as a marketplace that responds to diverse public demands, with the political parties viewed as the means for satisfying individual needs. Individualistic states favor economic development and reflect the belief that government should not intervene in the private sector, family, or church. Marshall, Mitchell and Wirt (1989) confirmed that political culture is an important variable in state education policymaking but can be constrained by the national policy agenda.

Louis and various colleagues have published extensively on the influence of political culture on education policymaking. Their work has demonstrated that political

culture impacts policy development and implementation, both in the U.S. and internationally. Based on their reviews of the literature on political culture, particularly the work of political scientists, Louis and colleagues utilize seven dimensions of political culture to analyze states' education policy behaviors, activities, and outcomes (Louis, Gordon, Meath & Thomas, 2009; Devos, et al., 2012; Louis, 2012):

1. *Openness*: Broad participation as opposed to dominance by the elite; multiple pathways for involvement of citizens and stakeholders;
2. *Decentralism*: Distribution of power, where decision-making rests closer to the site of impact;
3. *Rationalism*: Comprehensive approaches to policymaking, as opposed to unrelated initiatives or limited involvement of government;
4. *Egalitarianism*: Policy emphasis on the redistribution of resources or the minimization of disparities, where equity is upheld as an important value;
5. *Efficiency*: An emphasis on improving the use of resources for optimum policy impact;
6. *Quality*: Seeking effectiveness through state monitoring and oversight of public services;
7. *Choice*: Providing options to citizens regarding public services.

A study by Louis, Gordon, Meath, and Thomas (2009) examines U.S. political culture by analyzing accountability and standards policies in three states, two of which are the subjects of this dissertation: Missouri and Oregon. (The third state examined by Louis, et al. was New Mexico.) Overall, the authors find that political culture continues to

discriminate among the states, as illustrated by Elazar nearly half a century ago. Regarding Missouri and Oregon specifically, the study's findings support the conclusion of many previous empirical studies that Missouri has an individualistic political culture, while Oregon's political culture is moralistic.

The study illustrates that in Missouri, control of education policymaking resides at the executive level, with the governor, state legislature, State Board of Education (SBE), and Department of Elementary and Secondary Education (DESE) having the most influence. Organizations representing professional educators have a somewhat limited role in policy formulation, and once policies are implemented there is a history of limited oversight and enforcement from the state. By contrast, in Oregon many participants come together to discuss and formulate education policy. Collaboration among organizations is typical, and "groups, rather than individuals, are recognized as state policymakers" (p. 161). However, the researchers found that some inner circle and near circle players have a higher level of influence than others. At the time of the study, these groups included Oregon Department of Education (ODE), Oregon School Boards Association (OSBA), Coalition of Oregon School Administrators (COSA), Oregon Education Association (OEA), Oregon Business Council (OBC), and the Chalkboard Project, a coalition of five philanthropic foundations. Despite their dissimilarities, both states exhibit a historical preference for local control, giving local policy actors considerable power over the interpretation and implementation of education policy.

As part of a large study on school leadership, Louis, Thomas, Gordon, and Febey examine the effect of political culture on recent education policymaking (2008). The

authors explore the process of policymaking related to school improvement and accountability and how states' political cultures and policymaking histories influence their use of the four policy levers originally outlined by McDonnell and Elmore (1987): mandates – rules intended to produce compliance; inducements – money in return for certain actions, often accompanied by rules; capacity-building – money for investment in material, intellectual, or human resources; and system change – transferring official authority in order to alter the system by which public goods and services are delivered. The authors find that the use of particular policy levers is not directly linked to political culture, a finding that was unexpected. However, the study shows that political cultures are enduring, and they help explain the variation in states' approaches to school reform.

In another article based on the same broad study, Louis et al. (2010) investigate how differences in state political culture influence the way small school districts respond to state education policy. This question is particularly salient to this dissertation, which explores policy interpretation in small districts. The authors point out that research on small school districts is limited, with most studies of school reform efforts being based on medium to large districts. Further, the authors indicate that, because large urban districts are complex organizations unto themselves, they are less responsive to a state's political culture than small districts. The researchers studied four small districts, one in each of four states with contrasting political cultures, including Missouri and Oregon. (The other states were Texas and Nebraska.) The authors found that the districts responded to state education policies in a way that aligned to the assumptions inherent in their respective political cultures, and suggest that political culture directly determines how small districts

respond to policy initiatives. They state: "We may hypothesize that in traditional states, districts are more likely to see themselves as compliant actors; in individualistic states they view themselves as somewhat free to interpret standards in their own ways, and in moralistic states they are likely to see the states as partners in improvement" (p. 362).

Professional development as policy. Since the 1980s states have increased their efforts to influence classrooms, and policies related to professional development are on the rise. However, the link between such policies and increased learning is unclear (Louis et al., 2008). Recent federal policies have included requirements for professional development as well. The Obama administration's Race to the Top (RTTT) program stipulated that states receiving funding must support teachers through effective professional development that is informed by data and includes coaching and induction (U.S. Department of Education, 2015c), while the ESEA Flexibility Waiver program requires that teacher and principal evaluation results and student data systems be used to inform professional development (U.S. Department of Education, 2015b).

Professional development is a central goal of current state teacher evaluation policies. In sixteen states, teacher evaluation legislation explicitly identifies professional growth as the purpose of teacher evaluation, while 42 states have guidance documents expressing the same (Minnici, 2014). In *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement* (2013), Darling-Hammond lays out a vision of teacher evaluation as part of a system of supports for strong teaching throughout the profession. Darling-Hammond advocates that this teaching and learning system must address continuous improvement of individual practice and at the same time attend to

professional learning to improve practice among teams of teachers and within the profession as a whole. The author argues that most teachers experience a disjointed series of signals about professional practice during their preparation in college, in their early years of teaching, and as they move through the myriad of state and local requirements for continuing licensure and staff development throughout their careers. Bringing coherence to this arc of learning is key to Darling-Hammond's vision of a profession centered on well-articulated standards that guide professional growth.

Knapp (2003) lays out a depiction of professional development as a policy instrument for improved teaching and learning. He argues that understanding professional development as a policy pathway is necessary for visualizing how professional development efforts can improve the functioning of school systems. Knapp illustrates a "chain of influences" that connects the actions of "upstream" policymakers with "downstream" effects on teacher and student learning (p. 116). Among the "upstream" influences are federal, state, and district policies, implementation of professional development strategies, and improvement strategies taking place in other locales. Those that occur "downstream" include teacher knowledge and beliefs, teacher actions, and student learning. Knapp explains that "the professional development pathway" can be complicated by the fact that leadership for professional development tends to be disbursed among individuals and organizational units. Also, the professional development pathway is one among many (sometimes competing) policy pathways aimed at improved teaching and learning, making it is necessary to consider adjacent pathways and broader reform efforts.

Knapp finds that policymakers can influence professional learning environments in schools. Further, he points out that professional development initiatives show better results when they are coherently linked to other reforms and that professional development policies are more effective when they help support administrators' ability to guide instructional improvement. Although Knapp's work is not well connected to state policy research, his articulation of professional development as a policy pathway is salient to this paper's explication of the formative purpose of teacher evaluation, which is emphasized in many state teacher evaluation policies.

Recent research on implementation: Policy interpretation and mediation. Policy research has evolved over time, particularly the study of implementation. Generally, the literature recognizes three generations of research on policy implementation. This section provides a brief summary of these phases in the history of implementation research and then addresses recent research on policy interpretation and on the individuals and organizations that mediate between policymakers and implementers: policy intermediaries. It should be noted that, although this study does not purport to investigate policy implementation, a brief summary of recent research in this field serves to delineate policy *implementation* from policy *interpretation*, the latter of which is the focus of this study.

Three generations of implementation research. The first generation of implementation research focused to a large degree on the difficulty of implementing public policy. As deLeon and deLeon (2002) explain, this early implementation research, made up mainly of case studies, described "the immense vale of troubles that lay between

the definition of a policy and its execution" (p. 469). Researchers identified barriers to implementation, which included political factors, cultural issues, and a lack of will and/or capacity on the part of implementers. A second generation of implementation studies attempted to explain why some policies succeeded and others failed. This research aimed to explain the gap between policy intent and what was truly implemented. The results of these studies illustrated the characteristics of strong and weak implementation and aided practitioners in understanding how to implement policy (Fowler, 2009; Honig, 2006; deLeon & deLeon, 2002). The third generation of scholarship on policy implementation addressed the growing complexity of public policy, which saw less emphasis on individual initiatives and more attention to large-scale restructuring and reform. During the third generation of implementation research, scholars broadened the scope of understanding about the individuals, organizations, and conditions that affect implementation (Fowler, 2009; Honig, 2006).

Policy interpretation: Implementers as learners. Recent research frames implementers as learners who do much more than execute policy mandates. Scholars illustrate that those responsible for implementing policy in the local setting are interpreters of policy, making meaning from policy, explaining it, and determining next steps (Ball, Maguire, Braun & Hoskins, 2011a, 2011b; Hill, 2000). In education, much of this research has focused on teachers as interpreters of policies related to classroom instruction (for example, Cohen & Barnes, 1993; Spillane, 1999). Spillane, Reiser & Reimer are credited with advancing this approach to implementation research with their 2002 explication of implementers as sense-makers who draw on prior knowledge and are

influenced by values and emotions as they interpret policy demands. As this research approach has expanded, it has been recognized that outsiders – individuals and organizations beyond the school walls – play a key role in policy interpretation (Levinson, Sutton & Winstead, 2009; Ball, et al., 2011a). Levinson et al. argue that the process of meaning making involves policymakers and all others associated with a policy who must "negotiate a complex field of meanings and understandings" throughout all phases of the policy process (p. 779).

Policy intermediaries: Operating in the middle. Due to a limited amount of research on those policy actors who function between policymakers and implementers, scholars have begun fairly recently to explore more deeply the work of policy intermediaries and to define their role. These scholars argue that historically, implementation research has focused only on actors with a formal role in the policy process – policymakers and implementers – paying little or no attention to those who operate in the middle (Hamann & Lane, 2004; Honig, 2004; Coburn, 2005; Hill, 2000). Further, little research has addressed the messages sent by policy intermediaries as they mediate between policymakers and local policy actors (Hill, 2003). The studies described here are among those that shed light on the role of policy intermediaries and that are particularly salient to this study of teacher evaluation policy.

In a 2004 article on intermediary organizations as actors in education policy implementation, Honig asserts that such organizations are becoming more prominent, yet little is known about how they function. Further, she argues that a lack of research on these organizations has led to a limited understanding of what defines intermediaries,

allowing a "motley collection of organizations" to be identified as intermediaries, from technical assistance providers, to professional development organizations, to reform organizations (p. 66). Thus, Honig offers her own definition of intermediaries:

> Intermediaries are organizations that occupy the space in between at least two other parties. Intermediary organizations primarily function to mediate or to manage change in both those parties. Intermediary organizations operate independently of these two parties and provide distinct value beyond what the parties alone would be able to develop or to amass by themselves. At the same time, intermediary organizations depend on those parties to perform their essential functions (p. 67).

Though Honig's definition does not specifically rule out governmental agents as potential intermediaries, her framing of the problem suggests that organizations such as state education agencies (SEAs) might be considered arms of the policymaking apparatus, thus excluding them from consideration as intermediary organizations.

Honig studied four organizations that assisted in implementing policies on school-community partnerships in the Oakland, California schools. The four organizations mediated between the school district office and partnership sites. Honig's study confirms that intermediary organizations are "distinct actors in education policy implementation" (p. 82) and that they add value to the implementation process. The study demonstrates that the functions of intermediaries are context specific, as are their abilities to perform these functions, based on realities such as policy mandates and capacity. Further, the study shows that intermediaries can expand their effectiveness through participation with

other groups in a "web of intermediary supports that build on the strength of each organization" (p. 83).

Offering a "learning perspective" (p. 266) on policy implementation, Hill (2003) argues that traditional theories explaining how policies are implemented ignore the process implementers go through as they develop an of understanding of new a policy. Hill introduces the term "implementation resources" to describe those individuals and organizations who assist street level bureaucrats in interpreting a policy and its demands. The author defines implementation resources as "individuals or organizations that can help implementing units learn about policy, best practices for doing policy, or professional reforms meant to change the character of services delivered to clients" (p. 269). Hill envisions many types of actors as potential implementation resources, including academics, consultants, foundations, and professional associations; however, she specifies that implementation resources are best understood as nongovernmental actors, explaining that by restricting the term this way she hopes to encourage research on nonstate actors in the policy process. Further explicating her conception of implementation resources, Hill describes these individuals and groups as the "connective tissue between policy as written in legislation or bureaucratic regulation and the implementing agent" (p. 270). Implementation resources assist implementers in understanding a policy's meaning, its requirements, and the skills or practices necessary for successful implementation.

Hill reports on an exploratory study of the effect of implementation resources on the enactment of a federal community policing policy in five police departments. The

study finds that implementation resources promoted the new policy by assisting implementers in understanding the policy itself and what it meant in their everyday practice. In addition, implementation resources helped implementers develop the skills and competencies required by the new policy. The author argues that, as nongovernmental actors, implementation resources may be considered more legitimate by front line implementers than would representatives of a state agency, making them an effective instrument to help ensure the successful implementation of a new policy.

Although many studies of policy intermediaries frame these individuals and organizations primarily as nongovernmental actors (Coburn, 2005; & McLaughlin, 2010: Mitra, Sanders & Perkins, 2010; DeBray, Scott, Lubienski & Jabbar, 2014) – as do Honig (2004) and Hill (2003) above – Hamman and Lane (2004) specifically studied whether SEAs, which are part of the governmental apparatus associated with education policy, can and do function as intermediary organizations. In a study exploring whether SEAs function as policy intermediaries, Hamann and Lane asserted that policy messages are interpreted throughout the process of policy development and implementation and, further, that interpretation occurs at all levels across the educational system, with policy intermediaries assisting implementing agents in understanding policy. The authors reviewed the literature on the role of intermediaries and illustrated that such individuals and organizations engage in negotiating, bargaining, reframing, and sometimes co-opting policy as they operate between policymakers and the front lines of implementation.

In their study of two SEAs mediating a federal education policy, Hamann and Lane argue that NCLB has offered more discretion to SEAs in policy implementation

than did previous ESEA acts, giving them more latitude to adapt federal policy demands. The study finds that, as policy intermediaries, SEA personnel go beyond interpretation to become "coauthors" of federal policy, adapting the policy to suit the state context and integrating new policy mandates with existing policies and programs (p. 447). In each of the two states that are the subject of their research, Hamann and Lane find that the SEA had clearly and substantially shaped the policy that was ultimately implemented. The fact that SEAs adapt federal policy is viewed by the authors as inevitable, since SEA personnel cannot be considered "blank slates"; naturally, they employ their professional knowledge and experience to make sense of the policy and determine the best path forward (p. 448).

Summary of Reviewed Literature

This review of the literature on teacher evaluation practice and policy demonstrates that the evaluation of teachers has evolved from a process focused on minimal quality control to an approach that bears high-stakes consequences for both individual teachers and the schools that employ them. Now teacher evaluation is being utilized as a powerful policy lever to enact a theory of action that improving teacher evaluation systems will improve student learning. Not only has policymaking around teacher evaluation increased, the policies have become more specific, with schools responding to mandates from states and incentives from the federal government that promote particular approaches to teacher evaluation, including requirements to incorporate student growth scores in evaluation systems.

The focus on teacher evaluation policy is part of a general increase in education policymaking that began in the 1980s. The momentum for policy development in teacher evaluation specifically has been spurred by a variety of factors, including criticisms of existing practices, research on effective teaching, advancements in measurement, and a narrow "public" definition of teacher effectiveness. Other factors, such as the reform narrative that continues to emphasize equity, accountability, and U.S. competitiveness, and the more recent "bad teacher" narrative have contributed as well.

The literature reviewed here also illustrates tension between the formative and summative purposes of teacher evaluation, setting up challenges for policymakers and local educators. While states profess that the formative purpose is central to their teacher evaluation mandates, the political rhetoric and the narrowed definition of effective teaching suggest a predominant focus on the summative purpose. The potential role conflict for the school principal is exasperated by the fact that the field of educational administration preaches the value of formative evaluation, while the strong public emphasis on summative evaluation and high-stakes consequences for teachers and schools requires a modification of principals' past practice.

This review of the literature makes it clear that local policy actors face a complex challenge as they endeavor to interpret state and federal teacher evaluation mandates and design local systems. As Weaver-Hightower's ecology metaphor suggests, the work local actors must undertake to learn about policy, negotiate its meaning, and seek aid from policy intermediaries is embedded within an environment where governments, policymakers, voters, the media, policy histories, political cultures, ideologies, and

societal conditions – and the interrelationships among these factors – all influence local actors' policy interpretations and the outcomes of their efforts.

Conceptual Framework

The conceptual framework underlying this study addresses the context, influences, and actors involved in the process of policy interpretation, a phase of the policy process that occurs after formal policy adoption and before full implementation. The framework incorporates the following concepts: the ecology metaphor for educational policy described by Weaver-Hightower (2008); Elazar's political culture typology (1970); and the conception of policy intermediaries as aids in the policy process, which has been articulated by Honig (2004), Hill (2003), and Hamann & Lane (2004). As Figure 1 illustrates, policy ecology is conceived of as encompassing the other concepts; it is the context in which policy activity occurs, and it impacts the other policy influences and actors. Both policy intermediaries and local policy actors are presented in the diagram as nested within the influence of political culture. A dotted line separates policy intermediaries and local policy actors; this is meant to suggest the potential for influence to flow in both directions between these groups of policy actors. The figure also includes two arrows labeled "influence," indicating that the potential to affect policy interpretation flows both vertically and horizontally, between and within groups of actors and environmental influences that make up the framework.

Policy ecology. Weaver-Hightower's (2008) ecology metaphor for educational policy illustrates how a complex ecological system both affects and is affected by a particular policy. In this metaphor, as in science, a system of relationships exists in which

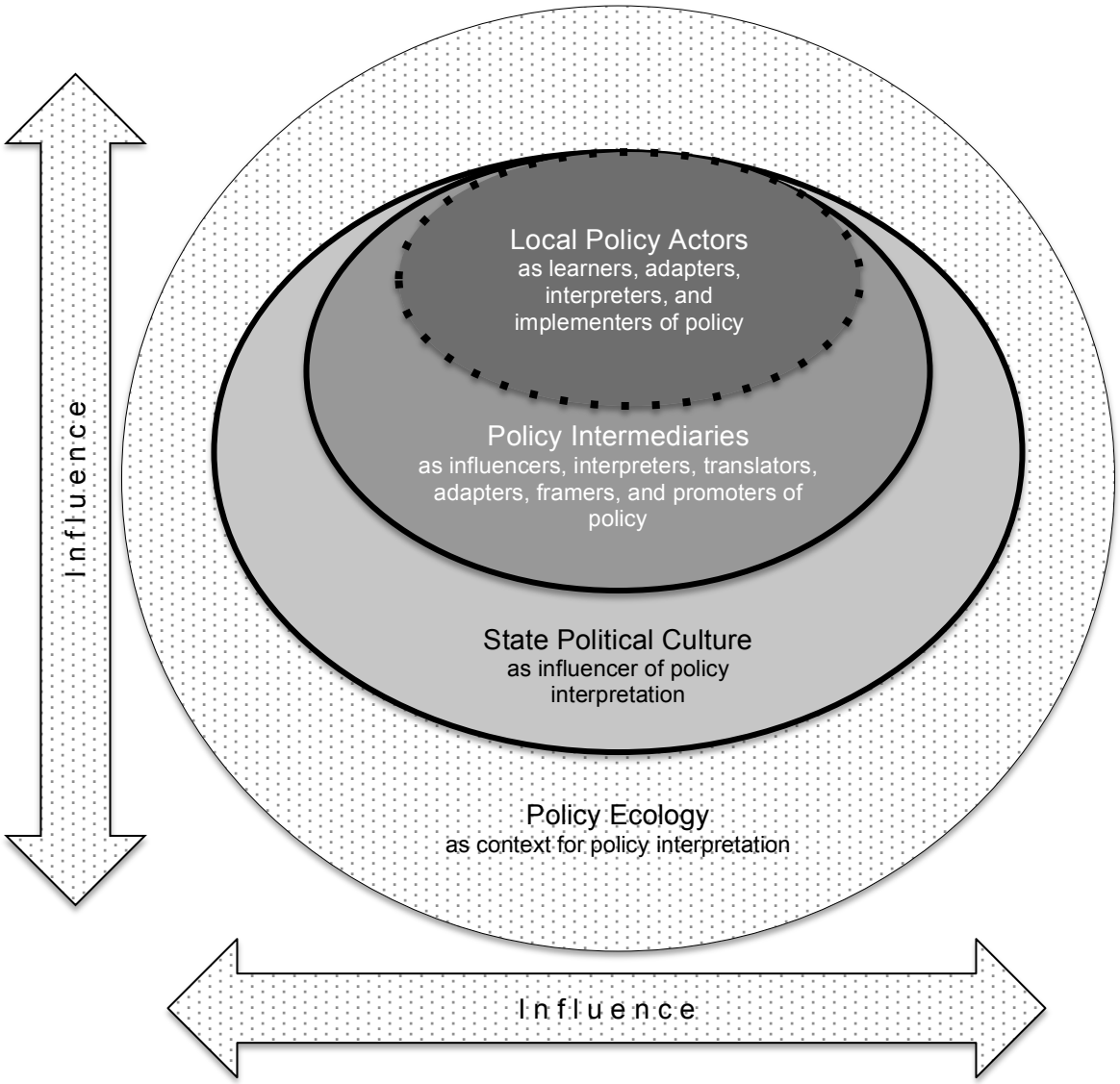

Figure 1. Conceptual framework for study: Interpreting teacher evaluation policies.

each factor in the environment influences the others, and "many complex interrelationships among them are required to sustain the system" (p. 155). In Weaver-Hightower's metaphor, the policy ecology consists of the policy itself, related policies, multiple actors, institutions, and organizations, as well as histories, traditions, and societal conditions. This conceptual framework envisions the policy ecology as a context

for the process of policy interpretation; it influences each of the other components of the process, political culture, policy intermediaries, and local policy actors, as depicted in Figure 1.

Political culture. Political culture is the enduring political attitudes and behaviors associated with groups that live in a defined geographical context (Elazar, 1970). Scholarly research has demonstrated that political culture impacts policy development and implementation and explains variation in states' approaches to school reform (Louis et al., 2009; Devos, et al., 2012; Louis, 2012; Louis et al., 2008). In Figure 1, political culture resides within the policy ecology; indeed, it is conceived of as a feature of this ecology. Further, political culture is depicted in Figure 1 as influencing both policy intermediaries and local policy actors. In addition, the figure supports the idea that political culture may affect the policy ecology.

Policy intermediaries. Policy intermediaries, those organizations and individuals who operate between policymakers and implementers, are distinct actors in the policy interpretation process. They add value by carrying out functions that neither policymakers nor implementers provide in a given situation (Honig, 2004; Hill, 2003). Policy intermediaries may or may not be associated with the various tiers of government involved in policymaking and implementation. For example, in some situations, SEAs, school districts, or individuals representing these groups may be considered policy intermediaries. Policy intermediaries perform a variety of roles, including but not limited to negotiating policy meaning, translating the policy, teaching about best practices to achieve effective implementation, promoting the policy's goals, framing or reframing the

policy and its expectations, and adapting the policy to the state or local context (Honig, 2004; Hamann & Lane, 2004; Hill, 2003). Figure 1 illustrates that policy intermediaries relate directly with local policy actors, that these two sets of actors influence one another, and that they are impacted by political culture and the broader policy ecology.

Local policy actors. In this conceptual framework, local policy actors are those individuals at the local level who are directly involved in discussions and decisions about the policy in question. Depending on the policy, the set of individuals identified as local policy actors is likely to change as differences in the topic of the policy and the nature of its requirements encourages individuals in different roles or with different areas of expertise and/or experience to participate. As represented by Figure 1, local policy actors are impacted by all of the other influences and actors in the process. In this conception of the policy interpretation process, local policy actors have the potential to influence policy intermediaries and to affect the broader policy ecology.

Policy interpretation. Traditionally, the task of interpreting policy – framed in this paper as a separate phase in the policy process referred to as "policy interpretation" – is considered to be part of the policy implementation process. However, this conceptual framework structures the actions associated with policy interpretation as distinct from implementation, necessarily occurring before a policy is put into effect and requiring deliberate attention and focused effort on the part of state and local policy actors. Since this research is not a study of policy implementation, the emphasis of this conceptual framework is on developing an understanding of how the teacher evaluation policy was interpreted during the time between its formal adoption and the date when it was required

to be implemented. This approach does not assume that policy intent is fully established in law and regulation, but rather is negotiated, adapted, and framed during the policy interpretation process. As policy research has evolved, scholars have identified the importance of interpretation in the policy process, consistent with the approach taken in this paper. For example, Honig (2006) argues that recent research on policy implementation frames interpretation is an inevitable aspect of the policy process; Ball, McGuire, Braun, and Hoskins (2011a) present interpretation of policy as a necessary process engaged in by local policy actors as they attempt to enact policy mandates; and Levinson, Sutton, and Winstead (2009) assert that the process of appropriation (which includes the activities I associate with interpretation) occurs "across and within the various institutional and microinstitutional sites where policy flows and takes shape" (p. 779).

Chapter Three

Methodology

This study was designed to explore the phenomenon of local policy actors – specifically school district administrators, building principals, and teachers – responding to new state teacher evaluation policies with regard to federal policy pressures and the national conversation about teacher effectiveness. The central research question addressed by this study is: How do local policy actors interpret teacher evaluation policies? Sub-questions include:

1. How do local policy actors interpret the *meaning* of teacher evaluation policies?
2. How do local policy actors experience the *process* of interpreting teacher evaluation policies?
3. How do state-level policy actors influence local responses to teacher evaluation policies?

The central research question can be addressed from two perspectives: *content* and *activity*, as indicated by the two sub-questions. In addition, the influence of state- level policy actors is pursued through the third sub-question.

The study's conceptual framework, articulated in Chapter Two, centers on local policy actors as learners who influence and are influenced by state policy intermediaries. In this framework, both sets of policy actors reside within and are affected by state political culture and the broader policy ecology. In this conception of the teacher evaluation policy interpretation process, influence flows between and within the groups of actors and environmental influences that make up the framework.

Research Design and Methods

A qualitative approach was chosen, since, according to Creswell (2014), qualitative research is useful for exploring the meaning ascribed to a particular problem by individuals or groups. In addition, qualitative research is helpful in tracing the process of meaning making (Merriam, 2009). Further, Creswell recommends qualitative methodology when little research has been done on an issue, so it is a suitable choice for this study of an emerging research problem: understanding the work of policy actors charged with designing local teacher evaluation systems in response to the dramatic growth in state and federal teacher evaluation policy. This research was conducted at a time when local school districts in nearly every U.S. state were simultaneously interpreting new state teacher evaluation policies and enacting local evaluation systems most in response to federal policy pressures.

This study employed a case study methodology, a fitting choice because of the opportunity it provides to investigate a contemporary phenomenon within its real-life context (Yin, 2014). According to Stark and Torrance, case study methodology is especially appropriate for studying new policies because a case study can be used to examine the policies "from the participants' point of view, as well as hold policy to account in terms of the complex realities of implementation and the unintended consequences of policy in action" (2005, p. 33). The study's design was an embedded, multi-case design, involving participants from multiple school districts as well as state-level stakeholders in two U.S. states. An advantage of multiple-case designs is that such studies can be considered more robust and the evidence produced more compelling than

studies with a single-case design. Further, analytic conclusions drawn from multiple-case studies can be considered more powerful than those from single-case studies (Yin, 2014).

Sampling and Data Collection

The two states that were selected for this research were Missouri and Oregon. These states were chosen from among nine states whose education policy histories have been studied extensively by Louis and colleagues. Some of Louis' recent research on this topic began with a large study on school leadership (Louis, Leithwood, Wahlstrom & Anderson, 2010). I had access to Louis and colleagues' data on education policy histories and on the political culture within this set of nine states. (The states were: Indiana, Missouri, North Carolina, Nebraska, New Jersey, New Mexico, New York, Oregon, and Texas.) Missouri and Oregon were selected because they are typical cases and can be viewed as representative of the larger population of cases of interest (Merriam, 2009; Seawright & Gerring, 2008). Oregon and Missouri are representative of the population of 43 U.S. states and the District of Columbia that have adopted new teacher evaluation policies as a requirement of their participation in the Obama administration's ESEA Flexibility Waiver program (U.S. Department of Education, 2015b). Additional relevant characteristics the two states hold in common include:

1. Neither Oregon nor Missouri is among the 19 states that have received federal Race to the Top (RTTT) grants to enact the Obama administration's favored reforms (U.S. Department of Education, 2014b); thus, teacher evaluation policies in the two states are a response to one federal policy, not two, and neither state received RTTT funding.

2. Policy timelines in both states required local implementation of new teacher evaluation systems during the 2014-2015 academic year. This timeline matches that of 16 other U.S. states (National Council on Teacher Quality, 2014b; Doore, B., Jorgensen, B., Doore, S., Mason, C.A., 2013).
3. Both states' teacher evaluation policies allow flexibility in the design of local teacher evaluation systems, where local districts may choose to enact a model designed or recommended by the state, design a local model, or adopt a model available publically or for purchase. According to Pennington (2014), of the 43 states and the District of Columbia that have ESEA Flexibility Waivers, 32 allow local discretion in system design. Consequently, the teacher evaluation policies in Oregon and Missouri can be considered typical of the majority of ESEA waiver states in terms of their approach to local flexibility around system design.
4. Both states are mid-sized states with few large cities, making them similar to one another in that neither state is among the nation's largest or smallest states, and in each state the number of major urban centers is limited to a few. Data from the 2010 U.S. census shows Oregon to have a population of approximately 3.83 million, ranking 27th among the 50 states, with three cities having populations over 150,000, while Missouri has a population of 5.99 million, ranking 18th among the states, also with three cities having populations over 150,000.

(For a summary of the status of U.S. states regarding ESEA waivers, RTTT, and teacher evaluation timelines, see Appendix A.)

Despite the features that make Oregon and Missouri similar to one another as typical cases for this study on teacher evaluation policy, the two states can be considered diverse from one another. According to Seawright and Gerring, the selection of diverse cases "has as its primary objective the achievement of maximum variance along relevant dimensions" (2008, p. 300). Relevant features on which Oregon and Missouri differ include:

1. The states have differing political cultures: Oregon is considered a moralistic state while Missouri is considered an individualistic state (Louis, Gordon, Meath & Thomas, 2009). Political culture has been shown to be an important variable in state education policymaking and helps to explain the variation in states' approaches to school reform (Marshall, Mitchell & Wirt, 1989; Louis, et al., 2008).
2. Although both states currently have Democratic governors, party control in the two state legislatures contrast; in Oregon both houses of the legislature are under Democratic control, while in Missouri both houses of the legislature are under Republican control. In addition, the states contrast in terms of Presidential politics; Obama won Oregon in both 2008 and 2012, while his opponent won Missouri in both elections (McCain in 2008 and Romney in 2012). According to Fowler (2009), since the 1980s, schools have increasingly become sites for ideological conflicts. Thus, it is useful to study states that contrast in terms of political control by the two major U. S. political parties, each of which reflects differing ideological orientations.

The units of analysis for the study were three small school districts in each state, a total of six school districts overall. The rationale for selecting small districts is aligned with that of Louis and Thomas et al. (2010), who chose small districts to study district leaders' interpretations of state policy. The authors argue that small school districts are understudied in the education policy literature, except in the area of school finance. In addition, the authors characterize medium and large districts as "complex organizational settings" that are more likely to drive policy than to be in a position to respond to it (p. 329). Further, although most U.S. students attend large school districts, most of the districts in the country are small. Hence, small districts can be viewed as representative of most school districts in the U.S. For this study, small school districts were defined as districts with student enrollments of 2,500 or lower. This is the same definition used by other researchers in studies where district size is a variable of interest (Louis & Leithwood et al., 2010; Melnick, 1987; Worner, 1981).

Districts chosen for this study were comprised of at least three school sites, also referred to as school buildings. This choice ensured that the task of local policy actors to interpret state teacher evaluation policy and make decisions about a district-wide approach bore a similar level of complexity – the task involved working across multiple schools, each with separate administrations and school cultures. To find lists of school districts by enrollment as well as information on the number of school sites in each district, I consulted data from the websites of the state education agencies (SEAs) in Missouri and Oregon. The total number of school districts that met the criteria – a student enrollment of less than 2,500 and at least three school sites – was 171 in Missouri and 56

in Oregon. In Missouri, this represents 31.7% of the state's 520 school districts, and in Oregon this represents 28.4% of the state's 197 school districts. (See Table 1.)

Table 1

Small school districts in Missouri and Oregon

	Total number of school districts	Number/percentage of small districts (those with fewer than 2,500 students)	Number/percentage of small districts with three or more school sites
Missouri	520	478/92	171/33
Oregon	197	140/71	56/28

Note. Data are from 2013-14.

Within this subset of school districts in each state, a purposeful sampling approach was taken in order to find the appropriate units of analysis (Merriam, 2009). That is, districts were chosen based on their approach to local teacher evaluation system design. In each state, three districts were chosen: one district that had adopted a teacher evaluation model designed or recommended by the SEA, one district that had developed its own teacher evaluation model or adapted one that was publically available, and one district that had worked with a direct support organization to develop a teacher evaluation model. For this paper, direct support organizations are defined as state-level groups that work one-on-one with school districts to support teacher evaluation efforts. I used publicly available data sources to determine which small districts in each state utilized each of the three approaches. These data sources included:

1. The websites of direct support organizations in the two states – the University of Missouri in Missouri and the Chalkboard Project in Oregon[5] (Curators of the University of Missouri, 2015b; Chalkboard Project, 2015b). (In the case of the University of Missouri, a phone call with a representative of the program was necessary to aid me in understanding the context for the data on participating districts.)
2. The websites of school districts in each state that met study criteria (student enrollments under 2,500 with three or more school sites).

In addition, I contacted the SEA and a state teachers association in each state for information on which approach to teacher evaluation system design each district in the respective states was taking. In neither state was a comprehensive list available from the SEA. It happened that a state teachers association in each state was able to provide some information about the approaches of various districts, and this information was used as a starting point in building lists of districts to contact for the study. District websites, as mentioned above, were used to supplement the information.

Letters or emails were sent to the superintendent in each of the districts identified as potential sites for the research, inviting the participation of the district. The letter explained that interviews were sought with district personnel who had been in involved in discussions and decisions about teacher evaluation, and a follow-up phone call from the researcher to discuss the study was promised (the letter is included in Appendix B.) In most cases, follow-up emails were also sent in order to solicit a response. The study was

[5] Prior to data collection and analysis, it was verified that the two organizations listed here were the only state-level direct support organizations providing services related to teacher evaluation system design in the selected states.

described to the superintendent in the follow-up phone call, and the superintendent was asked to participate and, if agreement to participate was secured, to provide the names and email addresses of study participants. After contacting many superintendents in each state and securing the participation of only one school district, it was apparent that offering an incentive for participation might improve the chances of getting superintendents to agree to participate in the study, and the offer of a $250 donation to a foundation that supported the local schools was made when contacting each succeeding superintendent. This proved to be a worthy method for securing participation from the required number of school districts.[6] Ultimately, the researcher contacted 22 school district superintendents in Missouri and 44 in Oregon before getting the requisite six superintendents to agree to participate in the study, three in each state.

When a superintendent agreed to participate in the study, he or she was sent a consent information sheet to share with potential participants (see Appendix C). It was requested that four to six individuals participate, including at least one district administrator, one building principal, and two teachers. Interviewing four to six individuals from each school district was an attempt to achieve saturation in data collection, and the decision to interview district administrators, building principals, and teachers was a purposeful attempt to get representation from the various internal stakeholder groups (Creswell, 2014). Next, the superintendent or a designee shared the names and contact information of study participants with the researcher. A total of 28 participants were identified, and one dropped out. Thus, 27 individuals participated,

[6] Despite the fact that the first school district had signed on without the incentive, I did contribute $250 to that school district's foundation.

including seven district administrators, seven principals, and 13 teachers. (See Table 2 for characteristics of Phase I study participants.)

Semi-structured telephone interviews with local teacher evaluation policy actors were conducted. Semi-structured interviews were chosen because of their strength for obtaining interviewee descriptions of their lived experience with and interpretations of

Table 2

Characteristics of Phase I study participants

	Number of participants	Sex		Avg yrs experience in education	Building level of principals and teachers[a]		
		Males	Females		H.S.	M.S.	Elem.
Missouri							
District administrators	3	2	1	18.3			
Principals	4	3	1	15.0	1	1	2
Teachers[a]	7	0	7	17.3	2	1	3
Missouri total	14	5	9	16.9	3	2	5
Oregon							
District administrators	4	3	1	25.3			
Principals	3	1	2	20.3	1	0	2
Teachers	6	1	5	22.3	2	1	3
Oregon total	13	5	8	22.6	3	1	5
Overall total	27	10	17	19.5	6	3	10

[a]One of the teachers was a teacher on special assignment and, thus, is not represented in the building level data.

the phenomena under study (Brinkman & Kvale, 2015). The selection of the telephone as an interview medium was made because of its advantages in terms of geographical coverage, time flexibility, and cost savings (Glogowska, Young & Lockyer, 2010; Wilson & Edwards, 2001). The interview protocol was designed to answer the study's research questions and to address the then draft conceptual framework. Broadly, interview topics included the state teacher evaluation policy, the process of decision-making on a local system, external sources of support, and the design of the local system (see Phase I interview protocol, Appendix D). Interviews were recorded and transcribed by the researcher. Two individuals from the study's first phase completed the interview questions over email due to scheduling difficulties. Interviews with the 27 Phase I participants were conducted between October 19, 2014, and January 29, 2015. These interviews lasted an average of 38 minutes.

To investigate prominent themes and uncover relevant contextual information about the perceptions described during the Phase I telephone interviews, a second phase of the study was conducted in which state-level policy actors were interviewed. Upon completion of the analysis of Phase I data (see Date Management and Analysis section below), Phase II participants were chosen using a snowball sampling approach, a common form of purposeful sampling where participants are chosen based on information provided by other participants (Merriam, 2009). As described below, the organizations that Phase II participants represented were selected based on Phase I data. After selecting four organizations in each state, participants were chosen based on their

job titles, available at the websites of the respective organizations.[7] A letter was sent inviting participation (see letter to Phase II participants, Appendix E). In addition, to verify that I had indeed selected the most appropriate representative of each organization, the first Phase II participant in each state was asked during his or her interview to recommend individuals from the identified organizations who would be good persons to speak to about teacher evaluation in the state. Recommendations aligned with the choices that had already been made. Of the eight individuals who were initially invited to participate in the study, seven agreed to participate themselves, and one recommended a colleague.[8] (See Table 3 for characteristics of Phase II participants.) Before scheduling the interview, each participant was sent a list of broad topic areas that would be addressed, along with a consent information sheet (see Appendix F).[9]

Table 3

Characteristics of Phase II study participants

	Representing				Sex		Avg yrs experience in current role
	SEA	Admin org	Teacher org	Direct support org	M	F	
MO	1	1	1	1	4	0	6.0
OR	1	1	1	1	1	3	4.9
Total	2	2	2	2	5	3	5.4

[7] My dissertation committee recommended that, ideally, four to six organizations in each state would be solicited to participate in Phase II. It was coincidence that ultimately the same number of organizations – four – was chosen in each state.
[8] The colleague was recommended because the person initially contacted was on medical leave.
[9] One Phase II participant requested to see all of the interview questions, so the interview protocol was provided.

As in the study's first phase, semi-structured telephone interviews of Phase II participants were conducted. An interview protocol was designed to investigate some of the same topics explored during Phase I, this time from a state-level perspective, and also to pursue themes that had arisen in the first phase of the study. In addition to the topics explored in Phase I interviews, Phase II participants were asked about the work of their organizations to support the teacher evaluation policy, policy history, and challenges to policy interpretation at the local level. (See Phase II interview protocol, Appendix G.) Interviews with Phase II participants were conducted between March 6, 2015, and May 5, 2015. These interviews lasted an average of 56.5 minutes. All but one of the interviews was recorded and transcribed.[10]

Data Management and Analysis

Data from Phase I interviews were coded via an open coding approach, consistent with grounded theory (Brinkmann & Kvale, 2015). This inductive approach was chosen because it allows the experiences of the interviewees to inform the data analysis, as opposed to using coding categories predetermined by the researcher (Yin, 2014). As new codes were added, already-coded transcripts were re-examined and amended as appropriate. During data analysis, I "winnowed" the data in order to identify the set of themes that later drove the design of the Phase II interviews (Creswell, 2014, p. 195). At the completion of each set of district interviews, transcripts and coding were reviewed, and a memo was written, capturing the district's story and noting key impressions (Orcher, 2005). This helped me to track the unique experiences and perceptions of

[10] Scheduling an interview with one Phase II participant proved to be particularly difficult, and ultimately, the participant offered an interview time when I was away from my recording device. Thus, I took handwritten notes on this interview.

participants in each of the six school district cases. NVivo, a type of computer-assisted data analysis software, was used to assist in coding and categorizing the data. Data were de-identified and stored on my computer, with names of participants stored in hard copy form, separate from interview data. The data were secured and were accessible only to me. The same process was followed with the data from the Phase II interviews.

Once all Phase I interviews were coded and memos about each school district written, two steps were taken:

1. A list of key state-level stakeholder groups was assembled, based on the data regarding the various individuals and organizations on which participating school districts relied for support and resources about the teacher evaluation policy. This list formed the basis for Phase II participant selection.
2. Themes were drawn out and utilized in structuring the Phase II interview protocol.

The process of analyzing Phase II data mirrored the process of analyzing Phase I data, with one exception: Due to the length and complexity of Phase II interviews, memos were created after each interview. It should be noted that, because this was an embedded multi-case study design, data from Phase I were analyzed in terms of their relationship to the pertinent three district cases *and* in terms of their relationship to the relevant state case, with all codes and categories analyzed at both the local and state (or summary) levels. Although there was some overlap in coding, data sets from the two phases of the study were kept separate in NVivo. Once all Phase II interviews were coded and memos written, themes were noted, and each theme was analyzed to determine its

prominence in relation to others. In addition, comments contributing to each theme were analyzed by organization and were compared to themes that had arisen during Phase I. At the conclusion of the analysis of Phase II interviews, all of the data together were analyzed by participant role (e.g., teacher, SEA representative), by district, and by state. Whenever possible, factual information was triangulated with data from succeeding interviews, documents analysis, or news stories (Merriam, 2009). Finally, a cross-case analysis was made between Missouri and Oregon wherein similarities and differences were examined, key findings from the literature review were referenced, and the research questions and conceptual framework were interrogated.

Throughout the process of analyzing data from both Phase I and Phase II interviews, codes and categories were enumerated (Orcher, 2005). In order to protect anonymity and to avoid making an inaccurate attribution of precision to a qualitative study with a non-random sample, in the results chapters that follow, the reporting of the numbers of respondents who expressed a particular view has been avoided. Instead, to communicate the relative prevalence of some ideas, non-numeric indefinite terms have been used to give the reader a sense of the prevalence of certain perspectives. Table 4 defines the terms used in this paper. It is acknowledged that there is some overlap in these definitions; this is intentional because it serves the purpose of allowing the tone and intensity of a set of comments to influence the term that is used. An exception to this practice was made when describing some of the Phase II results. In each state, since there were only four Phase II participants, it was more forthright to include the number of participants who expressed some viewpoints; in addition, because the Phase II

Table 4

Meaning of non-numeric indefinite terms used in Chapters 4 and 5

Term	Meaning
Couple	One to three
Few	Three to five
Several	Five to eight
Many	More than several but less than most
Most	Almost all of the referenced group
All	All

participants represented stakeholder groups that advocate certain points of view, reporting their responses in a straightforward manner is desirable.

As mentioned above, a variety of documents were analyzed in both of the study's two phases. In the main, these documents were utilized to triangulate data from the interviews. There are two exceptions to this. One, documents describing district and state teacher evaluation systems, frameworks, and models were considered evidence of the enactment of the policy and were utilized to help determine the interpretations of policy actors regarding the teacher evaluation policy. Two, documents reflecting the history of the policy were utilized to tell the story of policy background and development. Throughout the process of document analysis, the documents themselves were categorized and stored for reference (Merriam, 2009). With the exception of documents specific to the school districts associated with the study, documents were cited and included among the study's references. (School district materials were not cited in order to preserve the anonymity of the participating districts.)

Chapter Four

Results: Missouri

This chapter includes background on education policy in Missouri, including the education policy context and a history of teacher evaluation policy in the state. Next, the three Missouri school district cases are explicated, and finally the state case is presented.

Missouri's Education Policy Context

Political culture: Executive power with local control. Research on state political cultures has consistently identified Missouri as an individualistic state under Elazar's formulation (Louis, et al, 2009). In individualistic states, politics is viewed as a marketplace, and it is believed that government should be restricted and kept out of private affairs. Economic interests prevail, and political parties are relied on to get things done (Elazar, 1970; Fowler, 2009). In a 2009 study, Louis, et al. examined Missouri's political culture, focusing on education policymaking. The authors found that decision-making about education policy is centered at the executive level – with the governor, legislature, State Board of Education (SBE), and the Department of Elementary and Secondary Education (DESE) – and that professional organizations have a limited role. In addition, the authors illustrated that, as one would expect in an individualistic state, Missouri educators have a strong preference for local control regarding the details of policy implementation, stating, "Missouri has remained committed to a philosophy of decentralism and a preference for local control" (p. 171).

Governance: A recent change at the top. Like most states, in Missouri a SBE has responsibility for supervising public education (Fowler, 2009). Eight citizens

appointed by the governor and confirmed by the state senate make up the SBE, which has authority to appoint the Commissioner of Education, the chief state school officer (CSSO) in Missouri. The members of Missouri's current SBE have served in the state legislature, on local school boards, as educators, and in the fields of law, public policy, and business. All of them are male (Missouri Department of Elementary and Secondary Education, 2015b).

Missouri's CSSO, the Commissioner of Education, directs the Department of Elementary and Secondary Education (DESE), the administrative arm of the State Board of Education. As the State Education Agency (SEA), DESE oversees all public education programming, administers state and federal education policies, evaluates educator preparation programs, and advises local school districts (Missouri Department of Elementary and Secondary Education, 2015e). The current Commissioner of Education, Margie Vandeven, was appointed in January 2015, succeeding Chris Nicastro whose five-year tenure was plagued by controversy (Reischman, 2014). Vandeven is a veteran Missouri educator with a decade of experience in various leadership roles at DESE (Missouri Department of Elementary and Secondary Education, 2015a). Vandeven's appointment was somewhat controversial since she was promoted from within DESE; unlike many recent Commissioners, Vandeven did not rise to her post through the ranks of the superintendency in Missouri, and her preK-12 experience was in private schools in the state of Maryland (Robertson, 2014; Beckman, 2015). Vandeven was hired at a salary of $185,904 (Riley, 2014).

Recent policy history: An emphasis on improving the state's education profile. Currently, DESE and SBE are committed to the state's "Top 10 by 20" effort, which the two agencies established jointly in 2009, under the leadership of then Commissioner Chris Nicastro. The effort aims to have Missouri's students achieve at a level that places the state among the top 10 in the nation by 2020. The goals of the initiative center on college and career readiness, kindergarten readiness, educator effectiveness, and improved efficiency and effectiveness at DESE. Measurable benchmarks have been established, and DESE's strategies for achieving them are approved annually by SBE. Regarding teacher evaluation, the initiative specifies that all teachers in the state will be deemed effective by 2020 under the state's new teacher evaluation policy (Missouri Department of Elementary and Secondary Education, 2015g). According to Louis et al. (2009), Missouri has a history of promoting policies aimed at teacher quality. Among them are policies designed to support and improve teacher preparation, professional development, and mentoring for new teachers. A commitment to professional development is illustrated by the requirement in the 1993 "Outstanding Schools Act" that districts allocate one percent of their funding to professional development activities, determined locally by a professional development committee that teachers have a hand in selecting. The law is still in place (Missouri Department of Elementary and Secondary Education, 2015f).

Stakeholder organizations: Little influence, and no unity among teacher groups. As mentioned above, professional organizations are not particularly influential in initiating and developing education policy in Missouri, and Louis et al. (2009) found that

business groups had a greater role in education policymaking than did education stakeholders. According to the DESE representative who participated in this study, the department regularly seeks input from the state-level organizations that represent administrators and teachers as well as the Missouri School Boards' Association (MSBA) and the Missouri Association of Rural Education (MARE). However, representatives of the other state-level organizations that participated in this study express a desire for more influence as well as resentments over what is seen as a pattern of top-down directives.

One potential reason for the lack of influence by teachers organizations, in particular, is the fact that Missouri has three state teachers associations that tend to compete rather than cooperate. Only about three-fourths of Missouri teachers are unionized, and teacher unions are less involved in politics here than they are in most other states (Winkler, Scull & Zeehandelaar, 2012). Missouri's three teachers organizations include the Missouri NEA (an affiliate of the National Education Association), AFT Missouri (an affiliate of the American Federation of Teachers), and the Missouri State Teachers Association (MSTA), which is not associated with a national organization. Until 1972, MSTA was affiliated with the NEA, but due to disagreements about policies of the national affiliate, MSTA severed its relationship with NEA and became an independent association. A year later, Missouri NEA was founded as the state's NEA affiliate (Missouri State Teachers Association, 2015; Estes, K., n.d.).

According to the teachers association representative interviewed for this study, MSTA, the largest teachers organization in the state, is technically not a union; it is considered a non-profit educational organization and qualifies for federal tax exemption

under Section 501(c)(6) of the federal tax code (Missouri State Teachers Association, 2013b). Missouri's other teachers associations – Missouri NEA and Missouri AFT – are considered unions, filing as tax-exempt labor organizations under Section 501(c)(5). The representative asserted that part of MSTA's appeal is the fact that it is not a union. Missouri AFT represents teachers in St. Louis and Kansas City, while both MSTA and Missouri NEA have affiliates across the state, with most Missourians considering MSTA to be the rural teachers association and Missouri NEA to be the mid-sized city or suburban teachers association. According to participants in this study, this may not technically be true, but it is a perspective most Missourians share.

Complicating the picture further is the fact that the state does not have a collective bargaining law, and unions do not have the right to strike. Thus, Missouri is considered a weak union state (Freeman & Han, 2013). For the past several years, Republican lawmakers in Missouri have hoped to pass right to work legislation, which would make unions even weaker, and during the 2015 legislative session, Governor Jay Nixon vetoed a right to work bill (Yokley, 2015). About half of U.S. states have right to work laws, which give states the authority to determine whether employees can be required to join a union to get or keep a job (National Conference of State Legislatures, 2015). Generally, such laws prohibit unions from automatically charging dues to members of an employee bargaining unit.

Education outcomes: Average performance overall. In general, Missouri performs at or below average on most indicators of education quality. (See Appendix H for Missouri and Oregon education data in relation to the U.S.) Per pupil spending in

Missouri is more than $1,000 below the national figure, placing the state 30th among U.S. states. In Education Week's 2015 *Quality Counts* report, the state received a grade of C- and a rank of 33rd (Editorial Projects in Education, 2015b). The report addresses three broad categories of indicators: K-12 achievement, school finance, and chance for success (parent and family factors, early education, and indicators of preschool through adult participation and success). Missouri scored well on enrollment in full-day kindergarten, high school graduation rates, and steady adult employment. Areas where the state ranked poorly include advanced placement testing outcomes, and gains in National Assessment of Educational Progress (NAEP) results. On the 2015 *Kids Count* report, a measure of child well-being published by the Annie E. Casey Foundation, Missouri ranked 23rd on the education scale, which tracks four markers of success: preschool attendance, fourth grade reading proficiency, eighth grade math proficiency, and on-time high school graduation (Annie E. Casey Foundation, 2015). Compared to Oregon, Missouri has fewer students in poverty, fewer English learners, fewer nonwhite students, and a considerably higher graduation rate for students overall and for every racial or ethnic group (U.S. Department of Education, 2015a; Editorial Projects in Education, 2015a). Both states perform similarly to national averages on the National Assessment of Educational Progress (NAEP) (National Center for Education Statistics, 2015).

History of Teacher Evaluation Policy and Practice in Missouri

Teacher evaluation policy: An early focus on teacher improvement.

Missouri's original teacher evaluation law was passed in 1983, at a time when other states were enacting similar policies, partly in response to criticisms of teacher evaluation

practice in the country. Missouri's policy required school districts to conduct teacher evaluations, directing DESE to provide suggested procedures. Although the law required that teacher evaluations be conducted with "sufficient specificity and frequency," it was not mandated how often evaluations were to occur (Missouri Department of Elementary and Secondary Education, 2012, p. 1). In response to the new law, DESE issued a model teacher evaluation plan in 1984. This model was based exclusively on classroom observation and focused on teacher behaviors (Hughes, 2006), a common approach to teacher evaluation that ultimately drew criticism (Wise, et al., 1985; Ellett & Teddlie, 2003).

A 1990 study by Schweitzer found that over 90% of school districts in Missouri used the state's model, and most teachers preferred the system to the teacher evaluation practices they had experienced in the past, indicating that they felt the system had potential for improving teaching and learning. In 1999 the department released an updated version, called the Performance-Based Teacher Evaluation (PBTE) (Missouri Department of Elementary and Secondary Education, 2012). The new model was expanded to include clear evaluation procedures, a professional development focus, and an emphasis on training for both principals and teachers (Hughes, V.M., 2006). The PBTE was Missouri's model teacher evaluation plan until the teacher evaluation law was amended in 2011 and a new model system was approved in 2013.

Teacher evaluation practice: Compliance with little meaning. Although Missouri was an early adopter of teacher evaluation policy, the system was not viewed as meaningful for teacher improvement. Both local and state actors in this study commented

on the practice of teacher evaluation in Missouri before the current policy was developed, and it was clear that participants viewed teacher evaluation practice – and the policy requiring it – to have been inconsistent and even weak. In general, it seems that methods had not kept up with best practice. In addition, as implemented, there was a greater emphasis on summative over formative evaluation in spite of the stated intent of the policy. To state-level actors, it was especially important that the new teacher evaluation system bring consistency across the state, as standardization had dropped off over the years. According to the DESE representative, within the first several years after the PBTE had been developed, there was a "period of peak uniformity that then deteriorated" as districts made adjustments to the model to suit local needs and embed best practices. In addition, both local and state actors expressed concerns about the lack of training for evaluators.

Local actors offered specifics on weaknesses in teacher evaluation practice before the new policy was enacted. One teacher described the shift away from the former methodology to the new approach:

> The principal could come in and could basically pretty much say, that was a good lesson . . . There wasn't anything there about the depth of the teaching, and the follow through, and were these practices consistent?. . . I never felt like the old evaluation system was a true evaluation of what was going on in the classroom, where I feel like this one is because it's looking at the student data and it's looking at the practices that the teacher uses consistently.

A principal commented on the emphasis on professional growth contained in the new state policy and the value of increased observations in his local plan:

> I've been in this long enough – I remember when the principal would come in once a year and evaluate, and I don't know that that gives a true picture of what's really going on in a classroom. So this has caused us to do more evaluations and it's also caused the teachers to have to work on a growth plan.

However, despite these concerns about the particulars of teacher evaluation practice in Missouri, teacher evaluation *was* happening regularly before the new policy was adopted. Although evaluations were considered perfunctory, there was a general sense that school districts were complying with policy mandates and classroom observations were taking place regularly, though not frequently. Several of the teacher participants commented that they had received regular evaluations throughout their careers and that they were comfortable being observed. The representative of the University of Missouri described the situation this way:

> I'm not saying that there wasn't good teacher evaluation going on, but I don't think it was happening in a very consistent, widespread method in Missouri. It was very much a local call about what you're doing, how well you're doing it, and what you're getting out of it.

Development of the current policy: Fits and starts. Missouri's current teacher evaluation policy was created over a period of six years, and it is fair to say that, aside from a clear focus on local control, the process lacked continuity. In 2007, the Missouri Advisory Council of Certification for Educators (MACCE) initiated the development of

new teacher and leader standards in Missouri, with the intent of using the standards to guide the preparation, certification, and evaluation of teachers, principals, and superintendents. MACCE, which was established by the state legislature in 1985, is an advisory council composed of 25 members who are appointed by the SBE. At least 15 of the 25 members must be active classroom teachers (Mo.Rev.Stat. § 168.015.1). The work of developing the new standards that began in 2007 was led by DESE and involved over 100 educators representing school districts, colleges and universities, and a wide range of education organizations. According to the DESE representative interviewed for this study, the work on the standards got off to a good start, "stalled for a bit" and picked up again in 2009 (personal communication, April 29, 2015).

In 2010, the Missouri legislature passed SB 291, which required school districts to develop teacher and administrator standards and directed DESE to create model standards (Missouri Department of Elementary and Secondary Education, 2012). According to the representative of the teachers association interviewed for this study, SB 291 began as a bill related to virtual education. It became a broad-ranging education bill and among the amendments to the bill was the language about teaching standards that was a priority for one of the states' teachers associations (personal correspondence, May 2, 2015). The Missouri School Boards' Association had also included strengthening the state's educator evaluation system among its 2010 policy positions (Missouri School Boards' Association, 2010). An early draft of the standards language directed DESE to develop the teaching standards, which raised objections from some education stakeholder groups. Once the local control component was added, specifying that each district was to design

its own standards with DESE supplying a model, the other education groups dropped their opposition, and ultimately the bill passed with bipartisan support. This is not surprising in Missouri, an individualistic state that values local control highly. Aside from the opposition noted here, there was little if any involvement of educators and stakeholder groups in the development of the policy, supporting Louis et al.'s (2009) finding that in Missouri, involvement in education policy development is typically limited to legislators, SBE, and DESE. It should be noted that no funding accompanied the passage of SB 291, and according to the representative of DESE, there has not been any funding made available to districts since that time, neither from the legislature, from the federal government, nor from grants.

The DESE representative explained that upon passage of SB 291, the department conducted a crosswalk between the draft teaching standards under development and the requirements of the new legislation, and it was determined that the standards matched the expectations outlined in SB 291 (personal communication, April 29, 2015). The SBE officially approved the model standards in June 2011 (Missouri Department of Elementary and Secondary Education, 2012). These standards ultimately formed the basis of a new model teacher evaluation system. (For a summary of Missouri and Oregon's teacher evaluation policies and systems, see Appendix I.)

ESEA Flexibility Waiver: Few players at the table. In August 2011, U.S. Secretary of Education Arne Duncan announced initial plans to offer flexibility to states under the Elementary and Secondary Education Act (ESEA), and education leaders in Missouri expressed interest right away (Bock, J., 2011). Once the state made a decision to

move forward on applying for the ESEA Flexibility Program, an evaluation system design team was established, consisting of 10 individuals representing school districts, colleges and universities, and DESE staff. This small group was responsible for developing the state's model teacher evaluation plan, utilizing the already developed standards and aligning with Missouri statutory requirements and the educator evaluation requirements of the federal ESEA Flexibility Program. The University of Missouri representative who was interviewed for this study served on the design team and indicated that the work of developing the state model was rather rushed, occurring over a period of approximately four months. Members of the team convened regularly for intense, daylong work sessions.

According to the DESE representative, the department felt well positioned to develop a waiver application because of the state's work on teacher standards. He stated, "What we were telling people back then is, we're not starting new work to do the waiver; we're actually describing what we're doing and hoping it fits into the requirements of the waiver, which it did." Two state-level actors, the representatives of the University of Missouri and a principals association, criticized DESE for leaving the education community out the development of the waiver application itself. To the University of Missouri representative, this was typical of the department's handling of federal issues:

> They don't really do a great job of, well, what do you schools think of this? Or the superintendents organization, or the two major teacher groups. I'm not saying they don't care – they do . . . It's like we don't have to take the time to ask them because that will drag this out for six or eight weeks. We're just going to go ahead

> and write it, or we're going to go over here and get this finished so we can get back to Washington.

The representative of the principals association was critical of DESE's tendency to push "top-down reforms" without involving the state's educators. Regarding the waiver application, he exhibited some distrust of both DESE and the U.S. Department of Education (USDOE) by suggesting that DESE's decision to leave the education community out of the process was strategic:

> I think they were trying to meet the waiver of the law, but I think my own insight is that the department used that to promulgate more reform by saying, we've got to do this to meet the feds' needs. And the feds would be easy to say to the states, well, it's the state's decision . . . They could point fingers at each other but still could move some of the reforms.

Despite this criticism, all of the state-level participants praised DESE and the Commissioner of Education at the time, Chris Nicastro, for prevailing in negotiations with USDOE over the student growth requirement. The waiver guidelines call for data on student growth to be a "significant factor" in a teacher's evaluation, and many states address this requirement by assigning a specific percentage of a teacher's evaluation score to student growth measures. Missouri's proposal, although it emphasized that such measures be "significant," allowed local school districts to determine how to weight student growth among the multiple measures involved in evaluating each teacher. This was a big issue for the teachers association whose representative was interviewed for this study. His organization was the only education stakeholder group that did not support the

state's decision to participate in the waiver program. However, once the waiver was approved, the organization backed efforts to embed waiver requirements in local evaluation efforts. Regarding the student growth requirement, he explained:

> We fought really, really hard and we were successful, and I give the department a lot of credit for that. Probably the biggest fight they had with the feds, and probably the biggest fight the department had with us, was that on the using student data as part of the evaluation – it was no required percentage.

Missouri submitted its application for the federal waiver in February 2012, articulating the elements of the model teacher evaluation system that had been developed to that point and explaining the planned process for completing and piloting the model. Included in the waiver application were seven principles of effective evaluation, which served as the foundation for the model and as requirements for districts that would choose to take a separate approach. The seven principles are: use of research-based performance targets; differentiated levels of performance; probationary periods for new teachers and leaders; use of student performance measures; ongoing deliberate, and meaningful feedback; ongoing training for evaluators; and the use of evaluation results to inform personnel determinations, decisions and policy (U.S Department of Education, 2012). In addition to being included in the waiver, the principles were codified in the Missouri Code of State Regulations (2014). USDOE responded by asking the state to bolster its plan for including student growth as a significant factor in a teacher's evaluation (Yudin, 2012). After refining the waiver application and holding the negotiations on student

growth referenced above, the state resubmitted its application in June 2012 and received official confirmation of its approval that same month.

At the same time that the waiver application was being written, revised, and approved, Missouri continued its efforts to complete the design and development of its model teacher evaluation system. The SBE approved the state model for piloting in June 2012, and during 2012-13, the system was piloted in 105 of the state's 520 school districts. The pilot districts were diverse in terms of size, geographical location, and student population. Revisions to the evaluation system were made in the spring of 2013, and SBE officially adopted the model teacher evaluation system in May 2013 (Missouri Department of Elementary and Secondary Education (2013a).

A political distraction: The case of Amendment Three. In the spring of 2014, only months before the new teacher policy was to be fully implemented statewide, a ballot initiative threatened to derail years of teacher evaluation policy efforts and undermine local control over teacher evaluation methods. The initiative would have amended Missouri's state constitution to require evaluations of teachers based primarily on student test scores, to require that districts retain or dismiss teachers based on the evaluations, and to limit collective bargaining rights for teachers (Bergquist, 2014; Ballotpedia, 2015b). Amendment Three was sponsored by the Children's Education Council of Missouri (CECM), an advocacy organization promoting education reform, and it was supported by a large donation from a retired financial executive, a billionaire who is a prominent player in Missouri politics (Fang, 2014). The Amendment Three campaign was surrounded by controversy, including accusations of improper behavior on the part

of the Missouri Commissioner of Education, Chris Nicastro. About one year before Missouri voters faced the ballot measure, it was reported that the Commissioner had advised CECM, making suggestions for the wording of the ballot initiative in order to make it more attractive to voters (Ujifusa, 2013; Reischman, 2014). At the time, several legislators called for Nicastro's resignation (Shapiro, 2014). Other more high profile controversies plagued Commissioner Nicastro's tenure, and she ultimately resigned in November 2014.

Participants had quite a bit to say about the amendment and the issues it raised; their comments centered on two themes: the concern that the requirement in the amendment to measure school and teacher performance based primarily on standardized test scores is not a reliable method, making it unfair to tie such ratings to high stakes employment decisions; and the amendment as a politically charged effort to reform the education system due to a prevailing negative attitude toward public education. A principal expressed his view of how the amendment would have affected educators:

> If it would have passed it would have meant a very hard strain on our system . . . For myself as an administrator, I would have really been concerned if my evaluations were not thorough enough for keeping the right teachers to produce the test scores that were required in that amendment. And then you have the teachers' side of things; they would have been pressured to try to figure out how to motivate students to produce a test score that doesn't mean anything on their grade . . . You tie that in to an evaluation and a career move and pay, and all that

> to me is ludicrous . . . It's to me a political game all the way around . . . The sponsorship was to more privatize education; that's very difficult to do.

A teacher explained her concern about the impact the amendment's proposed reforms might have on the nature of teaching and learning in a school building:

> I would give up my tenure to keep them from doing what they're trying to do . . . to make it where it's competitive, where we're paid based on how our students perform . . . They would have made it where I'm competing with the other [teachers at my grade level] to see who is the best. And so therefore, teachers don't want to share. I don't want to give my colleague – well, I'm teaching this and it's working and I'm doing a great job so I'm not going to share it with you because I want to keep my job over you.

The education community in Missouri came out strongly against Amendment Three, with organizations representing teachers, school administrators, school boards, and parents uniting to defeat it (Protect Our Local Schools, 2014). Defeat it they did: 76% of voters voted against the amendment. In the county most favorable to the initiative, only 30% of voters supported the amendment (Boston, C., Kovacs, K., Alexander, M. & Sitter, P., 2014). Groups representing teachers, school administrators, school boards and others worked together to defeat the ballot measure. According to a state-level actor, the amendment "really united the education community in Missouri . . . [It] was defeated in every county in Missouri. It was huge." A district administrator also mentioned the unity that Amendment Three brought to the education community in Missouri:

> It's a rich millionaire that has got it on the ballot, and he's trying to take over education as we know it in Missouri, and it's pretty sad. But the educators realize that and are trying for one time to band together to get it stopped.

The representative of a principals association explained that educators succeeded in raising concerns about local control regarding the use of student testing in teacher evaluation: "[The public] wanted that to be a local decision; they didn't want that to come from the top. And I think that was a huge piece of defeating Amendment Three."

For the DESE representative, the defeat of Amendment Three and the unity of the education community against it could be explained by the collaboration of education stakeholders around teacher evaluation, which began several years earlier. He stated:

> One of the nice kind of side benefits is, again, here's this other work we've been doing – we've all invested in it. So that became really kind of the battle cry: Why do we need an Amendment Three when we've been working on this for years?

The current policy: Local design with elaborated principles. Missouri's current teacher evaluation policy is based on three sets of requirements: the 1983 law requiring evaluations of teachers; the requirement for districts to adopt teacher and administrator standards, established in 2010 via SB 291; and the seven principles of effective evaluation codified in the Missouri Code of State Regulations and articulated in the approved federal ESEA Flexibility waiver of 2012.[11] Under the policy, each Missouri school district must design or adopt a teacher evaluation system that addresses the seven essential principles of evaluation. If a district chooses to design its own teacher

---

[11] Missouri's policy encompasses both teacher and administrator evaluation, but in this paper, discussion is limited to teacher evaluation.

evaluation system or to adopt one provided by a third party, the plan must be submitted to DESE for review and approval (Missouri Code of State Regulations, 2014). Full implementation of the policy occurred in 2014-2015, but districts were given until 2015-2016 to fully embed the student growth requirement. According to data collected by DESE, approximately 110 Missouri school districts have adopted the state's model teacher evaluation system (personal correspondence, September 9, 2015). The department does not keep records on all of the various teacher evaluation models utilized by districts that have not adopted the state model. However, data from the University of Missouri indicate that 248 districts used its Network for Educator Effectiveness (NEE) system in 2014-15, and more districts have signed on to begin using the University's system in 2015-2016, bringing the total to 269 (personal correspondence, September 2, 2015). This represents just over half of the school districts in the state. The NEE system, described more fully later in the paper, is a research-based teacher evaluation model developed by the University of Missouri that meets the requirements of the state policy. It is made available to school districts for a cost.

Three Missouri School District Cases

As described in the Chapter Three, Phase I of this study consisted of interviews with local policy actors who were involved in discussions and decisions regarding the local approach to teacher evaluation. Three small Missouri districts comprised of at least three school sites were chosen, based on their approach to teacher evaluation system design. The three Missouri districts were: Byrne Creek School District, a district that adopted the state teacher evaluation model; Laurent School District, a district that

developed its own teacher evaluation model, and Flores Valley School District, a district that worked with a state-level direct support organization to develop a teacher evaluation model.[12] (For information on the characteristics of Phase I participants, see Table 2, Chapter Three, p. 91.)

Byrne Creek School District. In Byrne Creek School District, teacher evaluation efforts were tied to the district's work on implementing the Common Core State Standards (CCSS), and a single, large leadership team led both efforts. Educators in Byrne Creek were focused on raising student achievement scores in a district that has struggled on statewide proficiency exams. With limited input from teachers, and mainly because it was the superintendent's preference, the district adopted Missouri's state teacher evaluation model.

The community and its schools: Rural, low income, low-performing. Byrne Creek School District is located in a small rural community in Southeast Missouri where farming is an important part of the community's history and economy. Citizens are proud of the area's civil war history, and the town boasts several historical attractions. The median household income in Byrne Creek is significantly below the state average, and though the population of Byrne Creek is small, the district covers a large area – more than 350 square miles. The percentage of nonwhite students in Byrne Creek School District is highest among the three districts in this study, and more than two-thirds of students receive free or reduced-price lunch. (For data on community characteristics, district characteristics, district outcomes, and district educators for all three Missouri districts included in this study, see Appendix J.)

[12] Names of participating school districts are pseudonyms.

Byrne Creek School District failed to meet the state target in both English/language arts and mathematics in Missouri's Annual Measurable Objective (AMO)[13] calculation in 2014, and unlike the other Missouri districts that are part of this study, which were both within five percentage points of the target in each area, Byrne Creek School District fell below target by a significant margin – over 15% in English/language arts and over 25% in mathematics. Missouri's AMOs are based on a subset of the Missouri Assessment Program (MAP) exams, which include science and social studies in addition to English/language arts and mathematics; there are a total of 22 exams included in MAP. Byrne Creek performed poorly in comparison to state averages on MAP testing, the lowest of the Missouri districts in the study. Despite this, the four-year graduation rate for Byrne Creek students is far higher than the state average – in fact, it is the highest among the Missouri districts in this study – yet a smaller percentage of the district's graduates go on to post-secondary education than the state figure of 65%. The percentage of Byrne Creek educators who hold advanced degrees is the highest among the three Missouri districts in this study, while their salaries are below state averages. Most administrators in Byrne Creek School District have served in the district for at least five years, and the superintendent came decades ago as a first-year teacher and has never left.

Teacher evaluation decision-making: Ties to Common Core efforts. Byrne Creek School District assembled the teacher evaluation and CCSS leadership team in 2012-2013. A committee of approximately 20 educators, it included district

[13] Annual Measurable Objective is defined as the goal set by each state indicating the minimum percentage of students who must meet or exceed standards as measured by the state's achievement exams (U.S. Department of Education, 2015a).

administration, all building principals, and multiple teacher representatives from each school building in the district. The initial work of the leadership team centered on establishing teacher data teams in relation to CCSS efforts, and those teams were tied to the implementation of the student growth requirement in teacher evaluation. The district has a strong training emphasis on preparing teachers and administrators to develop student learning objectives (SLOs) to meet both CCSS and teacher evaluation requirements.

Decision-making regarding the choice of a teacher evaluation model for Byrne Creek School District occurred in the summer of 2013, when the full committee was not available. District office staff presented principals with various models to consider. Thus, teachers were left out of the decision on a teacher evaluation model, although they had membership on the committee and involvement in coordinating teacher evaluation and CCSS efforts. However, a teacher on special assignment, who worked at the district level, did participate in decision-making. The superintendent favored the state model, and ultimately this was the district's choice. It was clear from the interviews that it was the superintendent who pressed for the state model, with some participants feeling that the decision was heavy handed, and others expressing an understanding of the superintendent’s interest in aligning teacher evaluation work with other projects on which the district was collaborating with DESE. In addition, it seems that the state model was a safe choice in a puzzling policy landscape. As one participant stated, "I think we were in such a state of confusion – let's just go with the state model. I'm not saying that was necessarily the right decision; I'm saying that’s the decision we made." The model was

implemented early – in 2013-2014 – and during the time when interviews were conducted, planning was underway for the implementation of the student growth requirement.[14]

Attitudes about teacher evaluation work varied among interviewees, with some participants describing a collaborative spirit and others expressing frustration that decisions were made primarily by the superintendent with little input from others. When asked how district staff learned about the state policy and its requirements, one principal explained, "We were at a staff meeting with all the administration last year and basically our bosses handed it to us and said, 'Here, implement this.'" Despite such concerns, all participants were enthusiastic about local efforts to improve student achievement through both CCSS and teacher evaluation work. The teacher on special assignment mentioned above, whose responsibilities included peer coaching and coordination of the district's CCSS efforts, appears to have been at the center of discussions on teacher evaluation and played a key leadership role in implementation planning. In addition, this individual was responsible for the training of both principals and teachers on CCSS and teacher evaluation processes.

The density of policy collaboration on teacher evaluation was low to moderate in this district (see definitions of terms in Chapter One). Although a large representative committee provided oversight to teacher evaluation and CCSS efforts, decision-making was held onto by district administration. Further, it seems that committee members were excluded from some policy discussions, leaving them unaware of the state context and

[14] I learned in August 2015 that the district had dropped the state model. No details regarding the decision or the new approach to teacher evaluation were provided to me.

the details of policy requirements. In addition, reliance on external resources was limited, and in the main, teacher evaluation leaders were consumers of information from state-level stakeholder groups, with very little if any interaction taking place between local educators and outside organizations, except in training sessions. The leadership team in Byrne Creek referenced DESE's information on teacher evaluation, and they consulted with another Missouri district, one that was an early adopter of the state model; participants found this opportunity to hear from a district that had already implemented the model to be very valuable. Administrators mentioned having attended DESE training sessions at the regional professional development center, and to a limited degree they sought information from Missouri's school administrator organizations. Both teachers mentioned their reliance on MSTA publications and emails for information on the teacher evaluation policy; each praised MSTA as a trusted resource. However, as indicated above, outside the district office there was little awareness of the details of the state policy and its political context.

Participants viewed Amendment Three as a distraction to local teacher evaluation efforts. The amendment had been defeated a few weeks before interviews with Byrne Creek educators were held, yet there was a keen sense that concerns related to the amendment were still present. Participants described fear and frustration among their colleagues in the weeks leading up to the election, with a principal explaining, "Oh, there was very high concern. If you understand our student population – we have a very high rates of poverty – and there were several teachers, many, that were concerned." Discussions about the amendment's potential passage took over Student Learning

Objectives (SLO) training sessions in Byrne Creek, and it was reported by one participant that even regional teacher evaluation training sessions were disrupted by conversations about the amendment, saying, "It has been a constant, present fear." A teacher took on a dejected tone in describing her thoughts about the Amendment Three campaign:

> I feel like for some reason, overall a lot of people anymore look down on education, you know, that we're not doing enough . . . For some reason it's in a negative light, you know, to some extent. And you have those people who . . . will do anything against it because they don't think we should be tenured.

Perspectives on the policy: Accountability focus. Educators in Byrne Creek School District placed a strong emphasis on school accountability in explaining their understanding of the goals of Missouri's teacher evaluation policy and the reasons for its adoption. Documents related to the district's 2014-2015 implementation of the state model were not made available; however, after the district dropped the state plan and switched to a new model in Summer 2015, I was able to obtain documents describing the district's new teacher evaluation system, including its purpose and goals. These statements present a mixed bag: there is emphasis on both continuous improvement of instruction and accountability. However, likely because the district experiences considerable pressure as a low-performing district, in interviews, participants consistently cited low student achievement scores as an explanation for the policy's passage and emphasized data-driven instruction aimed at improved student learning as a key policy goal. When asked about her understanding of the motivations behind the policy, a teacher

said, “Oh, wow. Um, I’m sure part of it was partly the pressure from everyone of hearing how schools are performing.”

Like the others in this study, participants in Byrne Creek School District expressed significant frustration with policy guidance from DESE. There was irritation about having to wait for basic policy information as well as concern about guidance and training being delayed. A teacher stated:

> I feel like Missouri has had some bumps along the way because we've had some things come out of DESE that would then be changed. You know, like two months later. So it's been a little bumpy in my opinion, but I feel like we're in a better place now.

Despite their fears and frustrations concerning the state policy context, educators in Byrne Creek School District were surprisingly optimistic about the teacher evaluation policy and its implementation in their district. Increased opportunities for dialogue about teaching and learning, improved collaboration within buildings and across the district at large, and hope for the policy’s potential to impact student learning were among the things participants seemed most enthused about. Concerns about local implementation of the state model were limited to general concerns about the change process, fidelity to the policy’s intent, and fairness in relation to the student growth requirement. A teacher closed her interview with this optimistic comment:

> Even if it started out with a negative, I feel like luckily we've been able to turn it into a positive. And I feel like we are in a much better place – oh my goodness – we are in *such* a better place because of what the students are getting. We’ve

stepped up our game. So if we had to do it because of DESE, if we had to do it because of the Common Core, if we had to do it because of the new teacher evaluation system, at this point I'm so thankful it's happened, I don't even care.

Laurent School District. In Laurent School District, decision-making about teacher evaluation was dominated by district administration, with a lack of involvement from teachers. Without establishing a committee or team to study and make recommendations about teacher evaluation, district administrators sought input from principals and developed a local approach to teacher evaluation, based on a publically available model.

The community and its schools: High performing schools in a bedroom community. Laurent School District, set in the Southwestern part of the state, is located within 25 miles of one of Missouri's largest cities. The population of the town has grown significantly since 2000, yet citizens remain proud of the small town feel of their community. Nonwhite students make up less than 10% of the district's population, and the number of students who receive free or reduced-price lunch is considerably below the state figure of 50%. Annual per pupil expenditures in Laurent School District are lowest among the three Missouri districts in the study, more than $2,000 below the state average of $10,127. Laurent School District did not meet 2014 AMO targets in either English/language arts or mathematics, but the district was less than 5% below target in both areas, and in 2014, the percentage of Laurent students who scored proficient or better was above the state average on about three-quarters of MAP Exams.

Considerably fewer than half of Laurent School District's educators hold advanced degrees, well below the state figure of 58.9%, and the lowest of the Missouri districts in the study. Average salaries for both teachers and administrators fall significantly below the state average, while they are less experienced on average than educators in the other Missouri districts that are part of this study. The administration is stable, with all administrators having been in in the district at least five years. The current superintendent has been in place for nearly a decade.

Teacher evaluation decision-making: Dominated by district administration. Laurent School District began its efforts to develop a new teacher evaluation system before Missouri finalized the state model, with work starting on the local plan in 2012-13. (The state model was approved in June 2013.) The district administrator who was interviewed described the decision to develop a local model:

> We tried to look at the policy that was out there and make sure that we were doing something that was beneficial to [our district], and not just because everyone was doing it, or that somebody at the state was doing it, or it was the state model out there.

The local model was an adaptation of the work of Kim Marshall, an educational consultant who has written extensively on the topic of teacher evaluation and whose publically available evaluation rubrics have been adopted in many school districts across the U.S. Laurent's local plan was piloted in 2013-2014 and fully implemented in 2014-15. Like all districts in Missouri, Laurent School District did not implement the student growth requirement until fall 2016.

In Laurent, the density of policy collaboration on teacher evaluation was low. District administration took the lead in teacher evaluation efforts, making most of the decisions and discussing the teacher evaluation plan with principals at regularly scheduled administrative meetings; no teacher evaluation team or committee was put in place. Explaining the work of designing the local plan, one of the principal interviewees stated, "We have [several] administrators, so I would say at any given time, all of us had an opportunity to view the policy, make any adjustments as necessary, put the input into it." One teacher, the president of the local Community Teachers Association (Laurent CTA, an affiliate of MSTA), was looped in, but only over email. She served as the de facto liaison for all information to and from district teachers, and it is fair to say that input from teachers on the design of the teacher evaluation system was not sought by administration in any systematic or formal way. In describing the typical decision-making process in the district, the CTA president stated, "[Our superintendent] kind of comes up with it himself and then puts together a committee to kind of approve what he's done . . . He'll listen to our voices, but it doesn't necessarily mean it's going to change." It seems that the approach used to develop the local teacher evaluation system is the standard approach to decision-making in Laurent School District, and educators accept that this is the way things are done here.

For information and resources about Missouri's teacher evaluation policy, administrators from Laurent relied most on administrators from other school districts, gaining information through their professional associations and through regional meetings. No interactions with state-level stakeholder groups were reported; knowledge

of the policy was gained through publications and websites. Administrators referenced Kim Marshall's publications and reviewed other teacher evaluation systems, including the NEE system from the University of Missouri. Although she was not involved in decision-making about teacher evaluation, the CTA president stated that, as is her usual practice, she relied on MSTA publications for help in understanding the state policy, explaining, "They'll translate it for you so you can understand what's being said. So I kind of rely on that so I can understand exactly what the policy is, what the policy's saying." Overall, participants from Laurent School District seemed to lack a solid understanding of the state policy and its requirements. This was true even for the principals, who were involved in discussions about teacher evaluation. Further, some interviewees seemed to have a lack of knowledge of the specifics of the local plan.

Perspectives on the policy: Skepticism about policy's purpose. Study participants from Laurent expressed the opinion that Missouri's teacher evaluation policy was driven by non-educators interested in reforming education. Related to this belief was a frustration with the public's perception of the teaching profession and skepticism that the policy is linked to broader efforts in Missouri to end tenure protections for teachers. A district administrator noted his belief that the policy was motivated by politics: "I believe it was prompted by political means and very likely by, I guess, political donations." There was considerable concern among educators in Laurent about the state's plan for including student growth as a required measure in teacher evaluation. Interviewees seemed impatient for guidance from DESE about how to implement the requirement,

both in terms of logistics and in terms of making the requirement meaningful and fair in the local context. . One district administrator commented:

> Say we arbitrarily pick some type of test and say, well, the students aren't doing well on this test, and therefore, we need better teachers, when . . . there's certainly a lot more variables that go into student learning and growth and even performance for that matter. I like to think that our students' growth and education is represented by the tests that we choose, but that may not always be the case. There's different ways to assess students, and students show their learning different ways, not always on a paper pencil test.

Participants were not completely against the use of student growth data in teacher evaluation, however. As the district administrator stated, "Quite honestly I see student performance being part of teacher evaluation – I see that being a benefit if it can be shaped in the right way. That's a big if." Multiple interviewees expressed concern about the student growth requirement having the potential to negatively affect morale in the district; it is perceived that the requirement is likely to introduce competition among teachers.

When Laurent participants were asked about the goals of the teacher evaluation policy, responses were mixed; some focused on continuous improvement and others stressed accountability issues, including teacher quality.[15] One teacher put it this way: "I believe it is . . . for our kids to be ready for college and, you know, not to be falling

[15] As defined in Chapter One, the term teacher quality has to do with measurable teacher attributes and bears an accountability focus.

behind." Teacher evaluation documents provided by the district reveal a system focused on accountability and high expectations for teachers.

In articulating the strengths of the policy and the local plan, participants were focused on increased opportunities for observation of teaching practice and the resulting increase in feedback to teachers and constructive dialogue between teachers and principals. The new teacher evaluation system in Laurent represents a shift away from formal, hour-long observations occurring about once every three years – termed by one administrator, "a dog and pony show" – to a system of up to ten unannounced walk-through observations per year. Administrators expressed appreciation for the opportunity this brings to see "authentic learning in the classroom," while teachers were still getting used to the shift. One district administrator described his view of the most important aspects of the policy this way: "Looking at the student data to provide more, maybe, goals for the students and the teachers . . . set the bar a little higher." A principal explained teachers' reactions to the new system: "Teachers were hesitant at first, as they saw the new evaluation tool as more of a hindrance than as a system of support. The increased presence in the classroom took some time for teachers to become accustomed to."

Overall, interviewees were not optimistic about the state policy or local efforts. When asked whether teachers in the district are comfortable with the local plan, the teachers association representative replied:

> I think teachers are always concerned. I don't think we're ever at a point where we're comfortable. I hate to say that, but I think there's so much turmoil in

> education right now that I don't think teachers are ever comfortable . . . because of the way they're evaluating us, and I think that we have to – we feel constantly worried about those three little tests that they give every year.

The few positive comments about the locally designed plan focused on the increase in accountability for teachers and on the clarity the plan brings to expectations for teacher performance.

Flores Valley School District. Teacher evaluation efforts began in Flores Valley School District several years ago, before new state and federal policies mandated the work. After reviewing and experiencing multiple approaches to teacher evaluation through an open process, the district chose to adopt the University of Missouri's Network for Educator Effectiveness (NEE) system.

The community and its schools: A desirable place to teach in a rural part of the state. Flores Valley School District is located in a small city in rural South central Missouri where manufacturing companies are among the largest employers and the nearest major population center is more than 100 miles away. The largest of the Missouri communities included in the study, the city is home to a community college campus. The crime rate here is higher than both the Missouri and U.S. averages, and the median household income is significantly below the state average. Fewer than 10% of students in Flores Valley are nonwhite students, and over 60% of students qualify for free or reduced priced lunch, which is higher than the state figure of 50%. Like both of the other Missouri school districts in this study, Flores Valley did not meet 2014 AMO targets in either English/language arts or mathematics. However, the district was less than 5% below

target in both areas. On over two-thirds of the 2014 MAP exams, the percentage of Flores Valley students who scored proficient or better was above the state average.

Teachers and administrators in Flores Valley are well educated, with over 65% holding advanced degrees. In addition, several of the administrators in the district hold doctoral degrees. The average salary for administrators in Flores Valley is near the state average, while the average teacher salary is more than $2,000 below the state average. Nearly 100% of core courses in the Flores Valley schools are taught by "highly qualified"[16] teachers, higher than the other Missouri districts in the study. The superintendent has been in the district for over a decade, having served as a teacher and principal here before becoming superintendent. His contract was recently extended. Flores Valley is the largest school district within a two-hour radius. According to one participant, this results in the district being a desirable place to teach in a rural part of the state, "We kind of get our choice of the better teachers."

Teacher evaluation decision-making: An early start and broad participation. Flores Valley convened a teacher evaluation committee during the 2009-2010 school year in order to improve the existing evaluation system, which was considered weak by district administration. In addition, the district wanted to be ready for pending teacher evaluation policy changes. At the time, the teacher evaluation committee included about 20 individuals, reflecting a broad cross-section of district personnel. The committee included representatives from district administration, all building administrators, and

[16] NCLB requires states to measure and report the extent to which its teachers are "highly qualified." Teachers are considered highly qualified when they hold a bachelor's degree, full certification or licensure, and prove they possess content knowledge in each subject they teach (U.S. Department of Education, 2015c).

about a dozen teachers appointed by their respective building administrators. The school board adopted the new teacher evaluation system in 2011.

Overall, the density of policy collaboration on teacher evaluation was moderate in Flores Valley, with fairly high involvement among both administrators and teachers in discussions about teacher evaluation, as well as a commitment to study alternatives; however, involvement of teachers in the actual decision-making process regarding the local system was limited. Despite having recently implemented a new teacher evaluation system, in 2013-2014, Flores Valley was among the school districts and charter schools to pilot the state's model teacher evaluation system. During the same school year, Flores Valley's evaluation committee was reconvened in response to the new teacher evaluation mandate. Upon reconvening the committee, district officials issued multiple invitations for teachers to serve on the committee; every teacher was invited to participate, and the committee grew to about 40. Before assembling the reconstituted teacher evaluation committee, district and building administration reviewed several possible teacher evaluation models and narrowed it to two, limiting the choice for the committee to one of two options, the state teacher evaluation model and the model available from the University of Missouri. Informational meetings on the two models were held during 2013-14, and feedback from the entire teaching staff was collected and reviewed by the committee. Finally, in the spring of 2014, the NEE system was chosen. A teacher expressed satisfaction with the process: "We created a committee and we gathered responses from our faculty, and that's what went into choosing our evaluation system. It was not just an administrative decision; it was a whole district decision."

It was one of the assistant superintendents who led the work of the teacher evaluation committee, and interviews with participants indicated that he and the superintendent were trusted and considered strong leaders. Those who served on the committee felt that their voices were heard and that their input into decision-making was valued. Committee members relied on district administration for information about the state policy, and participants indicated that DESE and NEE were the primary external organizations whose documents they referenced in order to understand the policy and its requirements. Aside from district administration, no one in the district interacted directly with state-level stakeholders, with the exception of training sessions offered by NEE. In fact, it is fair to say that an understanding of the state policy, its context, and the related federal requirements was limited to a handful of administrators, with the remaining committee members concerned mainly about the advantages and disadvantages of the two models under consideration. One of the teachers did mention that MSTA had been a reliable source for information about teacher evaluation, while the district administrator who was interviewed cited the Missouri Association of School Administrators (MASA) as having provided the best documents about the policy.

Participants were unanimously supportive of the decision to adopt the NEE system, and they exhibited a deep understanding of the elements of the system, its goals, and its benefits. Consistently, participants mentioned the following advantages of the model: the immediacy of the feedback available to teachers, the professional development modules available online, and the intensive training for principals, which is perceived to have resulted in strong inter-rater reliability.

Perspectives on the policy: Fear and optimism mixed. Flores Valley had had a system of regular teacher evaluation for several years; thus, Missouri's new policy did not represent a huge shift for educators in this district, and it seemed that there was a culture of openness around administrator observation of teaching practice. According to one teacher:

> As a tenured teacher, by state standards I was only required to have an evaluation done once every three years. We always had principals in our classrooms more than that, but that was the required. [Before the new policy] I would say on average I might see a principal once per quarter.

Despite the fact that teachers were observed regularly in this district, teacher fear and skepticism about the new policy came up as a theme in all but one of the interviews and was more pronounced here than in other districts in the study. Teachers in Flores Valley believed that the policy intends to remove low-performing teachers using unreliable methods and that student academic performance will be tied unfairly to high stakes employment decisions. One teacher stated, "I think that there's a lot of fear – especially among the new teachers – about our worth being tied to a test." Regarding the local approach, because the NEE system includes a student perception survey, teachers were nervous about how data from this instrument would be used. Despite the teacher evaluation committee's efforts to involve the broader staff in discussions about the local teacher evaluation approach, fear and skepticism persisted. The principal participant stated, "The thing we constantly battle is teachers that don't look at this as a – they look at this as a threatening thing, and it's supposed to be non-threatening."

When asked about the goals of Missouri's policy, educators in Flores Valley most frequently cited objectives related to continuous improvement of instruction, while some mentioned increased student learning and success. A district administrator emphasized providing support to teachers, stating, "I think of it a lot like an IEP [individual improvement plan] – everybody with an IEP that's based on their individual needs." A teacher highlighted professional development resources: "To create some . . . ongoing professional development for our teachers . . . to provide us with tools, strategies and things to make our teaching better." A review of the district's teacher evaluation philosophy and belief statements indicates an emphasis on continuous professional growth for the individual teacher, best practices in teaching and learning, and a link to student development.

Interviewees perceived that the new policy was adopted in order to improve upon weaknesses in previous teacher evaluation practice in their state. Only a few participants mentioned the federal influence as a reason for Missouri's policy change, and there seemed to be limited interest in the political discussions surrounding teacher evaluation. However, most participants did articulate concerns about Amendment Three, which was on the November 4 ballot in Missouri, the election occurring about two weeks after interviews were held. Although the major backers of the amendment had ceased efforts to promote their cause, educators in Flores Valley still expressed anxiety over the potential that it might pass, and they expressed frustration about the objectives of those who endorsed the effort. Both administrator and teacher participants expressed mixed views on tenure (the amendment would have eliminated teacher tenure in Missouri), yet all

were unified in their concern about the proposed tie between student test results and personnel decision-making. The principal interviewee stated:

> I think that we're guarded because this prop three thing is going to tie a lot of things to how kids do on tests, and that's not the whole reason we're there. We're trying to raise an entire child . . . We're trying to create a total person that's going to be a productive member of society, and you cannot rate through performance on a test how well that mission is being accomplished.

Understanding of policy requirements seemed to be limited to district administration. Among the teachers interviewed, all of whom served on the district's teacher evaluation committee, there was a lack of understanding of the details of the state policy. It seemed that since both the state model and the NEE system met all policy requirements, the committee felt no need to explore details of the policy, focusing instead on choosing the right fit between the two options the administration had presented for consideration. As noted above, the district administrator interviewed for this study led Flores Valley's teacher evaluation work, and he expressed considerable frustration with shifting requirements and mixed messages from DESE. His complaints included conflicting information, changes in decisions about policy requirements, and lack of clarity about the required use of student growth scores. This comment captures the frustration:

> I would call the state department, and I would get guidance one way, and then within a month they would announce that they were going in a different direction.

So it was very difficult to try and make sure that I was being up front and honest with the . . . information that I was given.

Overall, participants from Flores Valley School District seemed positive about the policy and their decision to adopt the NEE system. As one teacher stated, "We're on the right path with what we've chosen."

Summary of the Three School District Cases

As described in Chapter Three, data from interviews with local policy actors were analyzed from both the local school district and state (summary) perspectives. This section provides results from the summary analysis of views of the 14 participants from the three Missouri school districts that were part of this study. This section highlights results as they relate to the research questions and conceptual framework.

Interpretations of the policy's meaning. The central research question of this study addresses how local actors interpret the teacher evaluation policy, while the first sub-question addresses the policy's meaning. Local actors' responses to certain questions reveal information about their interpretations of the policy's meaning, including queries about the goals of the policy, the reasons for its adoption, the most important requirements of the policy, and policy conflicts.

Local actors' understanding of policy goals: Emphasis on continuous improvement. Local policy actors in Missouri expressed the clear understanding that the main goal of the teacher evaluation policy is continuous improvement of instruction; yet issues of accountability were at the forefront for some. Overall, there was considerable consistency in local participants' perspectives on the goals of Missouri's policy, with all

14 local actors citing one or more of the same three policy goals, listed here in order of frequency: continuous improvement of instruction, improved student learning, and improved teacher quality. The first goal, continuous improvement of instruction, is directly aligned with the formative purpose of evaluation. All of the Missouri principals in the study cited continuous improvement of instruction as the policy's primary goal, while the responses of district administrators and teachers were balanced across the three goals.

Local actors who identified improved student learning as a central goal of the policy exhibited two different perspectives. The first group expressed the goal of improved student learning as an extension of the goal of continuous improvement, illustrating a link between continuous improvement of teaching and its potential effect on the learning of students. This approach aligns with the formative purpose of evaluation. Those who took the second approach expressed the goal of improved student learning in terms of accountability for student achievement results and meeting external policy demands, which aligns to the summative purpose.

Among the three Missouri school districts in this study, participants from Byrne Creek were least likely to cite continuous improvement of instruction as a key policy goal, with all participants from this district emphasizing summative issues when discussing the policy's goals. This is not surprising, given that Byrne Creek is a low-performing district where considerable time and resources have been devoted to raising scores on standardized tests. Under pressure to improve student achievement, this district has utilized both the teacher evaluation and CCSS policies to place a strong emphasis on

data-driven instruction and instructional goal-setting based on results of achievement testing.

Local actors' understanding of the reasons for the policy's adoption: Sensitivity to public pressures. Local actors in Missouri were less aware of the teacher evaluation policy's origins than its goals, and they viewed the policy in light of national and state reform and accountability efforts. There was less consistency in local participants' perceptions about why the policy was adopted than there was in their understanding of the policy's goals, and several participants expressed uncertainty about the motivations behind the policy. In general, participants centered on summative over formative issues, with most participants describing the reasons for the policy's adoption in terms of efforts to increase educational accountability or to advance a reform agenda. Additional reasons included improving upon the existing teacher evaluation system and meeting federal policy requirements.

The emphasis on educational accountability in local actors' perspectives on why Missouri adopted a new teacher evaluation policy may be linked to the Amendment Three campaign, which was viewed by local policy actors as a politically charged distraction to teacher evaluation efforts. Thirteen of the fourteen local actors were asked about Amendment Three, and many agreed that the amendment was a distraction to local educators as they were implementing new teacher evaluation systems.[17] It should be noted that interviews with participants from Flores Valley were held in mid- to late

[17] The Missouri participant whose interview was conducted over email was not asked about Amendment Three.

October, before election day, while interviews with participants from both of the other Missouri districts, Laurent and Byrne Creek, were conducted after the election. [18]

Local actors' views on the most important aspects of the policy: A focus on teacher growth. Teachers and administrators in Missouri were united in their belief that the most important requirements of the teacher evaluation policy addressed professional growth, with nearly all local actors mentioning requirements of the policy that contribute to continuous improvement of instruction. Examples include the value of the dialogue between the principal and the teacher, the professional development supports available as part of the teacher evaluation system, the focus on instructional practice, and the opportunity for more principal time in classrooms. Only a few local actors identified the student growth requirement as the most important aspect of the policy.

Policy conflict: Challenges at the local level. To understand more deeply how local actors interpreted the meaning of the teacher evaluation policy, the interview protocol included a question about policy conflict, which, as defined in Chapter One, is conflict between the goals of a policy, or between its goals and the reasons for its adoption. Several local actors identified policy conflict, and it was clear that these conflicts challenged the policy interpretation process.

A major theme in the literature on teacher evaluation is tension between the two main purposes of evaluation, formative and summative. When asked if they perceived policy conflict, some local actors articulated a conflict between the policy's formative goals and an agenda aimed at "getting teachers." Most of the individuals who noted this

---

[18] Analysis of interview data demonstrates that participants from Flores Valley were no more or less likely to identify the Amendment Three campaign as a distraction to local teacher evaluation efforts.

conflict were principals, and they indicated having difficulty convincing teachers that the policy's main intent is to help teachers improve their practice. This sensitivity on the part of principals to the tension between formative and summative evaluation may reflect the potential for role conflict to challenge those who evaluate teachers (Biddle, 1986; Campbell, D.J. & Lee, 1988; Stiggins & Bridgeford, 1985; Popham, 1988, 2013; Gmelch & Torelli, 1993; Clift et al., 1992).

Another area of concern was identified by a few local actors in response to the question about policy conflict: the use of student growth data to evaluate teachers. Participants identified conflicts inherent in implementing this requirement because of questions regarding the validity of using student achievement data, particularly data from standardized tests, to measure teacher effectiveness. In addition, participants reported conflict between the student growth requirement and the ability of local educators to implement it, particularly because they felt there had been a lack of clarity about what exactly was required at the school and classroom level.

Influence of state-level organizations on local policy interpretations. The second sub-question guiding this research had to do with the process local actors undertook to interpret the teacher evaluation policy, while the third sub-question concerns the influence of state-level policy actors on local policy activity. In addition, the conceptual framework depicts state-level stakeholders as having a potential influence on local interpretations of the teacher evaluation policy. To understand whether and how they were influenced by state-level stakeholder groups in their interpretation of the teacher evaluation policy, local actors were asked if the information and resources they

received from sources outside their school district affected their view of the policy. Their responses shed light on the influence of state-level policy actors.

As described later in this chapter, state-level actors in Missouri were intentional in sharing the message that the teacher evaluation policy was aimed at continuous improvement. Although this message seemed to take root in local actors' interpretations of the policy, as described above, they did not perceive state actors as having influenced their views on the policy itself. Several local participants responded affirmatively when asked if the information and resources they received from sources outside their own school districts had affected their views of Missouri's teacher evaluation policy; however, in their responses all of these individuals described the influence of external sources on their *attitude toward implementing the policy*, not their view of the policy's value or its intent. Participants indicated that shifting or unclear policy requirements from DESE negatively impacted their view of implementing the policy. However, after receiving clarity from other state-level stakeholder groups, participants felt more positive about the policy and about their districts' ability to implement it successfully.

State of Missouri Case

The second phase of this study involved interviews with representatives of agencies or organizations that had been identified by local actors as reliable sources of information and support on the state's teacher evaluation policy. (For information on the characteristics of Phase II participants, see Chapter Three, Table 3, p. 93.) This section summarizes results from the study's second phase related to the research questions and the conceptual framework, providing the interpretive lens of state-level policy actors.

Also, additional results regarding themes that arose over the course of the study, which were not specifically addressed by the study design, are described. These additional results include the perspectives of both local and state-level policy actors.

Interpretations of the policy's meaning. As stated above, this study's first research sub-question addresses interpretations of the teacher evaluation policy's meaning. The conceptual framework portrays state-level policy actors as influencers of local policy interpretation, but it also recognizes these actors as interpreters themselves who must seek to understand the teacher evaluation policy and its requirements in the context of state political culture and the broad policy ecology. Thus, in this section, we examine state-level actors' interpretations of the policy's meaning, including their responses to questions about the goals of the policy, the most important requirements of the policy, and policy conflict.

State-level actors' understanding of policy goals: A clear focus on continuous improvement of instruction. Interviews with state-level actors revealed that the organizations they represented viewed the formative purpose of evaluation as a centerpiece of Missouri's teacher evaluation policy; three of four state-level stakeholders in Missouri identified continuous improvement of instruction as the primary goal of the policy. The teachers association representative placed his remarks about the goals of the current policy in the context of how teacher evaluation had been approached in the past:

> I think the ultimate goal is to have an evaluation method in place that improves instruction. And I'll say before [the new policy] that I think it was just a matter of

> how you're doing. I don't think there was a push or a strive to actually use it as part of improvement for instruction.

When asked his opinion of the policy's goals, the University of Missouri representative said, "Teacher development, teacher growth – that would be how I would put it in just a few words." The DESE representative described teacher evaluation as one part of the state's broad approach to supporting teacher growth:

> The system from the beginning has always been about this development across the continuum of your career – preparation, getting ready, to early days of service, to ongoing days of service . . . That was always the intent behind it.

The fourth state-level participant, a representative of a principals association, viewed the main goal of the policy to be improved teacher quality. This goal is aligned to the summative purpose of evaluation, yet the participant linked this emphasis on teacher quality to a need for professional development supports for beginning teachers:

> I think they were trying to . . . help develop those teachers, provide a couple of years where they can have mentors, really help them, you know, get in there and really help them get better.

State-level actors indicated that it has taken time for Missouri educators to embrace the goal of continuous improvement, which represents a shift in the culture surrounding teacher evaluation, toward more emphasis on the formative purpose. One explained:

> In the first year . . . they really didn't understand what the goal was, that this is to be used to improve instruction. They were still looking on it as a model that may

> have evaluated the teachers as to whether or not they were going to reemploy them or put them on improvement plans . . . And now they realize that, yeah, we can use this to do that, but it's also used to improve instruction – or should be used to improve instruction.

State-level actors' views on the most important policy requirements: Mixed perspectives. There was little consistency in the responses of state-level participants to the question about the policy's most important requirements. Requirements that were mentioned by more than one of the four state-level stakeholders included the opportunity for local control in teacher evaluation system design and requirements related to continuous improvement of instruction, specifically providing support for beginning teachers and offering quality feedback to all teachers. Other policy requirements cited as important by state-level stakeholders were the requirement that evaluators be trained, the mandate that evaluations must be conducted annually, and the directive that districts must design a standards-based system centered on the state's seven principles of effective evaluation. Broadly speaking, state-level actors valued more the policy requirements related to teacher evaluation's formative purpose – professional growth – than they did the requirements related to its summative aims – employment decision-making and educational accountability, though some of the requirements they deemed to be most important had more to do with issues of design and procedure than they did the actual purposes behind the policy.

Policy conflict: Less perplexing at the state level. Most of Missouri's state-level policy actors denied perceiving conflict between the goals of the teacher evaluation

policy or between the policy's goals and the reasons for the its adoption. However, the one individual who did cite policy conflict noted the same two areas of concern that were addressed by local actors: formative vs. summative evaluation and the use of student growth data in teacher evaluation systems.

In citing a policy conflict between the formative and summative goals of teacher evaluation, the representative of a principals association explained that this issue is especially concerning for principals who possess fewer skills in managing the demands of teacher evaluation:

> You're trying to develop people, but at the same time you're trying to be critical of them, which mixes what I call growth and development with employment decisions. And so I think there's a natural conflict there . . . When you're trying to develop teachers, build good relationships with them and so forth, I think sometimes that you need critical review, you need good quality feedback, but then you have to step over that line into summative evaluation, which really is an employment decision. So because those two are connected, I think sometimes that can cause some issues for less skillful principals.

The same participant questioned the legitimacy of using student performance data in teacher evaluation when such an effort is problematic in high-needs schools, identifying this concern as a policy conflict as well. Not only did he express concerns about fairness between teachers and across schools and districts, he raised a concern about the ability of high-needs schools to attract and retain teachers under this policy:

> Anytime you start talking about student growth and student assessment data . . . it's definitely problematic, particularly with teachers and principals who are in needier schools, where poverty is a huge issue, parent involvement is a huge issue, all of those other factors that really weigh into student learning and student growth. And I – one of the fears from lots of educators is if there's too much emphasis put on student growth, student data, student performance on standardized tests, that it's going to be more difficult to keep people in those places that have the biggest needs.

Messaging and support from state-level organizations. This study's second research sub-question addresses the process of interpreting teacher evaluation policy, while the third sub-question speaks to state-level actors' influence on local interpretations. In addition, the conceptual framework focuses on the policy interpretation process and the potential influence of state-level policy actors. In Missouri, state-level education stakeholders were intentional in messaging their views of the teacher evaluation policy's goals, and these efforts did influence local policy activity. To understand how Missouri's teacher evaluation policy was framed by state-level stakeholder organizations, it is helpful to review their responses to the interview questions that addressed the messages their organizations attempted to communicate about the policy and their efforts to support local educators in the rollout of the teacher evaluation policy. Also included here is a discussion about collaborative efforts at the state level.

Key policy messages: Formative emphasis for most. State-level participants were asked directly about the most important policy messages they attempted to communicate about Missouri's teacher evaluation policy. Both the DESE and MSTA representatives indicated their organizations attempted to send the message that teacher evaluation must be centered on continuous improvement of instruction. The representative of the University of Missouri did not address the question about his organization's policy messaging in his interview. However, as stated earlier, he did articulate his belief that the central goal of the policy is continuous improvement of instruction. The University of Missouri's NEE website presents the following statement exemplifying the organization's emphasis on this policy goal:

> The fundamental premise of NEE is simple: There are opportunities for improvement in the professional practice of all educators. We assist districts in identifying those opportunities for improvement in each educator so districts respond to those opportunities to go beyond the minimum requirements of compliance regulation to provide comprehensive, effective, truly helpful solutions for teachers and schools; to do all those things in a way that generates meaningful data; and ultimately to mine that data and from it learn still more about effective teaching and learning (Curators of the University of Missouri, 2015).

The only one of the four state-level stakeholders who did not list continuous improvement of instruction among his organization's key policy messages was the representative of a principals association. This participant indicated that the three main messages his organization attempted to send about the policy were: the policy must be

"doable"; principals must be recognized as instructional leaders and have other responsibilities taken off their plates in order to make the policy a success; and the decision about how to structure the student growth requirement within the teacher evaluation system must be made at the local level. Regarding the first two messages, the participant stated:

> It has to be something that they can wrap their arms around and feel like they can do a good job with it. And I think time is a huge factor in it, whether it really can be done . . . We have to take some of the other stuff off the plate.

Research going back as far as the mid-1980s has addressed concerns about time demands for evaluators as a complicating factor in the implementation of teacher evaluation policies (Stiggins & Bridgeford, 1985).

Support and resources provided by state-level organizations. State-level education stakeholders in Missouri engaged in a range of efforts to support the new policy, yet in the main, DESE acted on its own to lead the development of policy guidance, training, and support. This is to be expected, since DESE, as Missouri's state education agency (SEA), is charged with leading schools in the implementation of both state and federal education policies. Moreover, as previously mentioned, historically policy work in Missouri is restricted to inner circle players, primarily the legislature, the SBE, and DESE. For the other organizations that participated in this study – the University of Missouri, a teachers association, and a principals association – assisting local school districts in meeting the requirements of Missouri's teacher evaluation policy was viewed as a choice.

Missouri Department of Elementary and Secondary Education: Solely in charge of rollout. In its efforts to disseminate information about the teacher evaluation policy, DESE provided two broad types of resources to Missouri schools and districts. The first is the Educator Growth Toolbox, which is an online repository of resources. This repository includes policy documents, research, guidance on the teacher evaluation requirements, videos, webinars, and samples of materials such as observation forms and training guides, (Missouri Department of Elementary and Secondary Education, 2015c). According to the DESE representative, all of the materials housed in the Educator Growth Toolbox have been created within the past five years and these materials represent everything districts need to implement the policy. He views this repository as a helpful set of tools for individuals who are "self-initiated" with the motivation to seek out resources on their own.

A second category of support provided to Missouri schools was described by the DESE representative as, "the reach out." To execute this category of support, DESE utilized personnel in the 11 regional professional development centers (RPDCs) to provide training and materials on teacher evaluation in their respective areas of the state. Each of the RPDCs serves certain districts, and they offered training on a variety of teacher evaluation topics as well as providing materials to districts in their area. The RPDCs were established in Missouri about 20 years ago to provide access to professional development and policy support in all regions of the state. Although RPDC staff are not employees of DESE, the support and training they provide is based on the priorities of the department, which are specified in a contract between DESE and each of the RPDCs. In

addition, the department provides a considerable amount of funding to support the RPDCs (personal correspondence, April 29, 2014). The DESE representative described the department's work with the RPDCs as a way to help roll out new policy mandates and to ensure that all school districts in the state get the information and documents they need, stating "They reach out to those districts and say – here's what's going on, here's the training we've got, here's the materials we have – just to make sure that everyone gets the word.".

In addition to the RPDCs, the department utilized its area supervisors – employees of DESE who are housed in offices across the state – to monitor districts, track their participation in teacher evaluation training sessions, and offer help as needed. The DESE representative stated, "We just don't want anybody not knowing, so we'll knock on doors, make phone calls, whatever, to make sure they're getting themselves moving." Additionally, the DESE representative noted the difficulty of getting information to each of Missouri's 520 school districts:

> It feels like we have talked about this non-stop for five years with anybody who we could get in the room and listen. And so you tend to feel like . . . you've told everybody. But there is a whole lot more bodies than you realize. And so you really haven't talked to everybody. You've talked to a bunch of people, but there's a bunch a bunch a bunch of people! . . . So messaging is a challenge just because I think people underestimate the scope of what scale really means. Scale is a challenge.

In addition to concerns about scale, the DESE representative acknowledged another factor complicating the rollout: the readiness of local educators to receive policy messaging and guidance, especially with a policy as complex as teacher evaluation.

Teachers association: Ambiguous policy position. The teachers association that participated in this study was not an active player in the rollout of the policy; however, the organization was seen as a reliable source of information about the policy. In addition to stressing the goal of continuous improvement, the association also advocated that decisions about how to enact the policy's student growth requirement must be made at the local level. When asked about a communications strategy, the participant described a reactive, rather than a proactive approach. He explained that decisions about topics to address in communications with the organization's membership were driven by questions received by the association's leadership and staff, indicating that choices about what information needed to be shared were based on "anecdotal" information. The organization also ensured that their field staff – employees who work directly with members in offices across the state – were kept abreast of the latest developments on the policy and on the organization's guidance about how to assist local educators as they planned for implementation.

In addition, the teachers association posted information and guidance about the policy on its website, including materials on available teacher evaluation models, DESE resources and information on training sessions, and details about Missouri's ESEA waiver. These resources also included the organization's policy position on teacher evaluation. Among its 2013 policy positions, the organization advocated teacher

evaluation as a means of continuous improvement, promoted training for evaluators, pressed for involvement of teachers in local decision-making about teacher evaluation, articulated the need for clear standards for teacher improvement plans, and expressed opposition to the use of student achievement data as a sole criterion in evaluating teachers as well as opposition to the release of evaluation results to state or federal agencies.

Principals association: Focused on "doability." The efforts of the principals association that participated in this study may be seen as member-focused implementation support. Like the other professional associations in the state, the organization was not instrumental in system-wide efforts, and direct interactions with local policy actors were limited. As stated above, the representative explained that his organization focused on the system being manageable for school principals. According to this participant, the organization's primary means of supporting implementation of the teacher evaluation policy was through a collaborative project designed to promote effective teacher evaluation. Titled "Missouri Supporting Educator Evaluation" (MOSEE), the effort was sponsored by DESE, the Missouri Association of Elementary School Principals, the Missouri Association of Secondary School Principals, and the Missouri Professors of Educational Administration. MOSEE took place during the 2014-2015 school year and was conducted in two phases: the first phase involved two-day training sessions with an evaluation expert, while the second phase consisted of a series of "community of practice" sessions held in various locations across the state. MOSEE was aimed at ensuring quality observation and feedback in teacher evaluation (Missouri Department of Education, 2015d).

Despite the fact that the representative of the principals association did not include continuous improvement of instruction when listing his organization's key policy messages, his comments about MOSEE demonstrate that a central theme of this initiative was helping teachers improve their practice. He stated:

> [The training] focuses more on the positive side as opposed to what I call the hammer side . . . knowing full well that you kind of have to do something with those that are not doing – but if you encourage them and develop them along the way through positive feedback, it's a better chance of helping them improve their teaching.

The program's sponsors wish to continue the MOSEE program, but this will depend on funding (personal correspondence, May 5, 2015).

University of Missouri: A continuing role in supporting teacher policy. Whereas the University had played a lead role in designing Missouri's teacher evaluation model in the 1990s, now the organization stepped forward again, not only to serve in state policy planning, but also to provide a teacher evaluation system for district to purchase. The University's system, dubbed the Network for Educator Effectiveness (NEE), is one of the products and services offered by the Assessment Resource Center (ARC), a unit within the University's College of Education. ARC assembled a team of researchers and K-12 practitioners who consulted with state officials and education stakeholders to create the NEE system. The research-based system, which meets the requirements of Missouri's teacher evaluation policy, includes observation tools, student surveys, a web-based data tool, and online professional development modules. Districts that purchase the NEE

system are provided with training and field support as well (Curators of the University of Missouri, 2015). Charges include an annual rate per evaluator – which buys access to the online system, staff orientation sessions, summer training sessions for evaluators, and support from University of Missouri staff throughout the year – and an annual rate per teacher for access to the online data system and educator training library (personal correspondence, April 15, 2015).

The University of Missouri representative explained that there were two motivating factors behind the University's decision to develop a teacher evaluation system. First, there was a precedent: the University's involvement in developing Missouri's Performance-Based Teacher Evaluation (PBTE) model; educators were again looking to the University for leadership on teacher evaluation. Secondly, some at the University felt the need to fill what they viewed as a void; it was perceived that DESE did not act quickly enough to provide a model teacher evaluation system after the passage of SB 291, leaving districts without an "answer" on how to move forward. Work began on the creation of the NEE system in the spring of 2011, and nine districts were selected to pilot the plan in the fall of 2011. The following year saw an additional 22 districts joining the network and utilizing the NEE model. Since that time, the number of NEE districts has grown dramatically, with over half of Missouri districts using the model.

Because the University had developed a teacher evaluation system that school districts could purchase, interview questions about communication strategy let to a discussion with the University of Missouri representative about how the NEE group marketed its system. It was explained that the nine school districts that piloted NEE in

2011-2012 were chosen based on personal relationships and geographical location; an attempt was made to represent the various regions of Missouri. Once the pilot was complete and it was time to seek additional districts, there was no marketing plan or budget. NEE relied on word-of-mouth to attract more districts to the system. The NEE representative explained:

> If you look at the map today you can kind of tell . . . wherever there was one of our original nine districts, there is a hotbed. There's 30 districts around that one. And I think some of it is, you know, five of my neighbors are doing it, why am I not? Or what have they got that they are playing with over there that I don't have?

Impact of state-level collaboration: Limited yet positively received. In Missouri, there was little reference to state-level collaboration in interviews with local policy actors, but the topic was included as a specific probe in state-level interviews. Such collaboration was limited in this state, but there is evidence that a growing spirit of cooperation exists among education stakeholder groups, and it is appreciated at the local level.

In his interview, the representative of DESE articulated his perspective that the department was very intentional in involving educators and education stakeholder groups in the development and rollout of Missouri's policy. He described the major education stakeholders in Missouri as a "little family" on whom the department relied for input on teacher evaluation. This family responds when called upon to weigh in on policy initiatives, and according to the DESE representative, this group includes all of the organizations representing superintendents, principals, and teachers as well as the

Missouri School Boards' Association (MSBA) and the Missouri Association of Rural Education (MARE). In response to the question about collaborating with other state-level stakeholders, the DESE representative stated, "We've just had partners in organizations and practitioners in the state all along for the ride on this, so there's really not been any part that we didn't collaborate on."

Both in the interview with the DESE representative and in supporting documents describing Missouri's teacher evaluation efforts, there was evidence that a large number of stakeholders – practitioners from school districts and institutions of higher education as well as representatives of state-level organizations – were consulted with in the development of the state's teacher and leader standards (Missouri Department of Education, 2013a). Yet despite DESE's enthusiasm about what the department believes to have been a highly collaborative process of designing and rolling out the new teacher evaluation system, local actors seemed unaware. Furthermore, the comments of the other state actors did not support DESE's perspective that there was considerable collaboration in the development and rollout of the teacher evaluation policy. Especially regarding the development of the ESEA waiver application, other participants were quite critical of DESE for leaving educators out of decision-making. It seems that during the period when DESE was working on the development of the teaching standards and related tools – before the passage of Senate Bill 291 (SB 291) – the department was quite intentional in involving stakeholders in the process. However, once the work on teacher evaluation became associated with federal mandates and time pressures intensified, DESE moved forward without the same degree of input from the field. Indeed, there was a marked

contrast in the number of people who participated in the development of the standards – upwards of 200 – and the number who served on the design team for the state's teacher evaluation model – only 10.

There are signs that a new spirit of collaboration is alive in Missouri, partly as a result of the cooperation that was required to defeat Amendment Three. As already noted, the education community came out strongly against the amendment, and even organizations that are traditionally opposed to one another (e.g., the associations representing school boards and teachers) united to get the message to voters that the amendment was not in the best interest of Missouri students. Multiple participants noted that the unity among educators during the amendment fight was unprecedented, and they expressed the belief that this unity was key to the amendment's defeat.

The principals association representative framed the Amendment Three fight as the second of two successful collaborations among state-level education stakeholders, which he viewed as new and unparalleled. According to this participant, the first came during the 2014 legislative session when Commissioner Chris Nicastro, needing the support of education stakeholders to pass her recommended education budget, gained the assistance of the principals associations, the teachers associations, and the school boards association by agreeing to negotiate on student testing. The stakeholder groups sought a reduction in the number of state mandated tests, and they were able to secure the Commissioner's agreement before publically supporting her budget. According to the principals association representative, "That was probably really the first big collaborative

effort that we had . . . and that was the first time we really came together."[19] In addition, the principals association representative described two more collaborations in which his organization has been involved most recently. The first is MOSEE, described earlier, and the second is a newly launched effort to build a leadership development continuum that articulates the professional development needs of principals from the phase of aspiring to move into school administration, through the certification process, and across the career. This collaboration involves both of the principals associations (elementary and secondary) along with DESE, representatives of higher education school leadership programs, and other stakeholders, and it is expected to continue in the coming years.

It seems that the principals associations' efforts to collaborate across stakeholder groups have paid off, at least in good will. When asked about the feedback received from educators about the collaborative efforts in which the principals associations have been involved, the representative indicated that educators have responded positively; principals in the field appreciate the fact that state-level groups are teaming to provide policy support:

> People like that a lot. They feel like, you know, educators are pretty powerful when we can come together and agree on a strategy or plan . . . My folks feel like we are at the table, and our voice is being heard. And it hasn't always been heard.

Additional Results

In addition to results related directly to the research questions and the conceptual framework, three themes emerged regarding the teacher evaluation policy and its impact. Each of these themes came up without prompting. First, local actors expressed

[19] There is a lack of supporting resources regarding this episode.

considerable frustration over shifts in policy requirements and guidance, termed policy shifting in this paper. Secondly, participants considered the task of interpreting the policy and planning for its implementation to be more challenging in certain types of districts. Finally, at both the local and state levels, some participants were notably optimistic about the teacher evaluation policy.

Policy shifting: Frustration and anger. A prominent theme in interviews with local actors was the shifting of policy requirements as the teacher evaluation policy was being rolled out, which caused considerable frustration at the local level. Participants from each Missouri school district registered complaints about this issue, indicating that DESE changed its guidance about the policy over time, that there was a lack of clarity on policy requirements – especially the student growth requirement – and that information was generally slow in coming.

The policy shifting about the student growth requirement had to do with Missouri's interactions with the U.S. Department of Education regarding the waiver, yet none of the local participants appeared to perceive this. Perhaps this lack of awareness of state-level policy activity around teacher evaluation is representative of the individualistic political culture in Missouri, which doesn’t place a high value on participation in the policy process, leading to little understanding of education policy initiatives at the local level. Whatever the reason for the lack of awareness on the part of local policy actors, they expressed considerable frustration with DESE – even anger – in describing the shifts in policy requirements and guidance.

Teacher evaluation considered more challenging for certain districts. Because this issue arose as a theme during the study's first phase, a question about this topic was included in interviews of state-level actors; participants were asked whether they felt the policy was more challenging to implement in small and/or rural school districts or in districts with a high percentage of students in poverty. None of the local policy actors in the Missouri school districts that are part of this study identified this issue. (Local participants in Oregon did.) However, two of four state-level actors in Missouri did express the belief that designing and implementing teacher evaluation processes is more difficult in certain types of districts.

The teacher's association representative expressed the belief that the work of meeting policy demands was most difficult in small school districts because these districts possess fewer resources. In addition, this participant stated that in districts with few administrators there could be issues with inter-rater reliability:

> In larger districts you are able to have principals go to other buildings and evaluate the teacher that that building [principal] has already evaluated. They can go back there [and say] "This is where I saw the strengths and weaknesses" . . . I don't know that you're getting that in the small districts.

The principals association representative indicated his perception that the tasks associated with meeting teacher evaluation policy demands are more challenging in both high poverty and rural districts. Regarding rural districts, he expressed a concern about the reality that there is a lack of training available to help evaluators learn the skills needed to evaluate teachers effectively.

Policy optimism: The new policy as a change for the better. The new teacher evaluation policy in Missouri finds both state and local policy actors optimistic. In expressing their optimism, some emphasized the perspective that the policy is a positive change, while others expressed hope about the policy's potential to improve teaching and learning in the schools.

Some local policy actors registered optimistic comments in response to the question, "What do you feel are the most promising aspects of your local plan as it evolves?" Others provided optimistic comments in various places throughout the interview, several during the closing when they were asked if they had anything to add that had not come up during the conversation.

Typical of the comments by local actors who focused on the teacher evaluation policy as a positive change is this comment by a district administrator:

> There's more collaboration with the teacher, between the teachers and the principals, and more planning, more goal setting. A whole lot more collaboration than there ever was before.

One principal stated simply, "It's been a good move. It's a step in the right direction."

Participants who expressed hopefulness about the teacher evaluation policy spoke of the future under the new policy and its potential to improve teacher evaluation as well as teaching practice. A teacher noted the potential for the new system to improve her teaching: "There are things out there then for my administrators to say, maybe this is something we need to look at . . . Something that could actually help me grow as a

teacher." Another teacher expressed enthusiasm about the policy's potential, explaining how the new system changed teachers' mindsets:

> I am so excited about what it's going to do, I think, for teacher competence because the teachers are now looking at – every single night – what's happening in their rooms. And they're looking at the data, and the data never lies . . . We have turned our district. We've come from a place where teachers did not want to see the data . . . There was always an excuse. "Well, but that child is poor, that child is Black . . . You don't know the parents." . . . Now the focus has now been placed, personally, where it always should have been placed, on the instruction.

Among the three Missouri school districts that participated in this study, educators in Flores Valley and Byrne Creek were far more likely to make optimistic statements, with every interviewee from Flores Valley expressing optimism. Only a single participant from Laurent expressed optimism about the policy, registering enthusiasm that the new system brings accountability to classrooms and schools. These results are not surprising. In both Flores Valley and Byrne Creek, district administration was intentional about involving principals and teachers in discussions about school improvement efforts. Flores Valley began its work on teacher evaluation early, before Missouri adopted a new policy, and every educator in the district received multiple invitations to participate in the process. In Byrne Creek, teacher evaluation efforts were tied directly to the district's work on CCSS, with educators across the district participating in school improvement efforts aimed at increasing student performance on achievement testing. On the other hand, in Laurent, decision-making about teacher evaluation was limited to administration,

fear and skepticism among teachers was high, and on top of being excluded from decision-making, teachers received no training on or orientation to the new teacher evaluation system.

In general, state-level policy actors expressed hope and confidence about Missouri's new teacher evaluation policy. Least optimistic was the principals association representative who indicated that the policy can meet its goals only if it is implemented with integrity. In addition, he expressed concerns about difficulties for principals at the local level: "I think the question is still out, whether the evaluation system as it's laid out is really doable, given all the other tasks the principal has to do."

Both the representatives of the teacher's association and the University of Missouri expressed considerable optimism about the policy's potential, with the University of Missouri representative stating, "I can see in our numbers, we're [already] impacting classrooms." The most optimistic of the four state-level stakeholders was the representative of DESE who expressed confidence that local educators will implement the policy effectively and will innovate as their work on teacher evaluation unfolds:

> As districts take it and make it fit their context and think about how they're going to adjust it and make it fit – that's when the innovation is really going to come about, and then really important things are going to happen. I don't have any doubt that it's going to go and do more than we ever envisioned . . . If you give them some direction and they understand why you think this is the right direction, and then say . . . Given the challenges of the learning of your kids, what do you

need to do to the system to make it work? And when it becomes about the kids, that's when I think it'll go way beyond how we envisioned it.

Chapter Five

Results: Oregon

This chapter includes background on education policy in Oregon, including the state's education policy context and a history of teacher evaluation policy in the state. The three Oregon school district cases are explicated next, and finally the state case is presented.

Oregon's Education Policy Context

Oregon's political culture: High participation in education policymaking. Consistently, scholars of state political culture have identified Oregon as a moralistic state (Louis, et al, 2009). In moralistic states, politics is viewed as contributing to the public good, and government is generally trusted. Widespread participation in politics is valued in moralistic cultures, and the public exchange of ideas is viewed as an important part of the political process (Elazar, 1970; Fowler, 2009). In their 2009 study of education policymaking in Oregon, Louis and colleagues found that policy discussions involve many participants and that collaboration across stakeholder groups is common. However, some inner circle and near circle groups in Oregon have more influence; at the time of the Louis, et al. study, these groups included the Oregon Education Association (OEA), the Confederation of Oregon School Administrators (COSA), the Oregon School Boards Association (OSBA), and the Oregon Business Council (OBC). Louis and colleagues also indicated that policy approaches that allow flexibility in design at the local level are preferred in Oregon. The authors explain that this preference has remained

strong, despite the fact that the state is "moving in the direction of increasing state mandates" (p. 158).

Governance: Recent changes to the status quo. Until recently, Oregon's education governance structure was fairly typical. Oregon was one of about a dozen states whose chief state school officer (CSSO) was elected, not appointed. This Superintendent of Public Instruction was responsible for leading the work of the Oregon Department of Education (ODE), which is Oregon's state education agency (SEA). In contrast with other states, however, Oregon's CSSO was the lowest-paid in the country, making an annual salary of only $72,000, less than top ODE employees who were paid over $100,000 (Melton, 2011a; Cole, 2012). A recent analysis illustrates dramatic inconsistencies in the compensation of elected officials and agency heads in Oregon. The analysis showed that most agency heads were paid more than elected officials, including the governor, whose salary was $98,600. In addition, agency size and the relative experience of its leader appeared to have no influence on compensation (Hoffman, 2014).

Changes came in 2011 when then Governor John Kitzhaber endorsed two bills that would restructure education governance in Oregon, part of his vision for improving and streamlining the education system. The first bill proposed making the Governor the Superintendent of Schools with the authority to appoint a deputy, who would serve as CSSO and lead the work of ODE. The bill was supported by the Oregon Business Association (OBA) and the state's school superintendents. In opposition were teachers associations and the Oregon School Boards Association (Melton, 2011c). The second new law, which was less controversial, established the Oregon Education Investment

Board (OEIB) and created the position of Chief Education Officer (CEO) to lead OEIB. OEIB was charged with bringing unity to all facets of education from birth to college and career, driving education policy, and making recommendations to the legislature regarding education funding priorities. The new law established that, in addition to chairing OEIB, the governor appointed its members. OEIB was designed to work in concert with the other education agencies in the state, including ODE. Under the new structure, both the Deputy Superintendent and the CEO reported directly to the Governor (Oregon Education Investment Board, 2015a; Oregon Secretary of State, 2015).

One thing that has not changed is the fact that a State Board of Education (SBE) oversees the pre-K-12 system as well as public community colleges in Oregon, though now the SBE works closely with OEIB. Appointed by the Governor and confirmed by the Senate, the current members of the SBE bring diversity of gender and ethnicity and represent a range of occupations, including law, tribal government, the private sector, and both pre-K-12 and higher education.

Although the changes made by the 2011 legislature have been in place for only a short time, the departure of Governor Kitzhaber and a series of resignations among high profile education leaders have caused some to question whether the reforms will hold. After only a year on the job, the state's first CEO, Rudy Crew, resigned to take a position as a college president. His high salary ($285,000) and his frequent out-of-state travels were difficult for some to swallow. Crew's successor, Nancy Golden, was paid roughly 80% of what Crew made (Hammond, 2013; Hoffman, 2014). Governor Kitzhaber resigned under fire in February 2015 amid accusations of influence peddling, and then

Secretary of State Kate Brown replaced him. Not long after Brown took office, the state's first appointed CSSO, Deputy Superintendent Rob Saxton, resigned, having served only three years but with a record of having shaken up the status quo (Hammond, 2015a). Brown then appointed the current Deputy Superintendent, Salam Noor to the post (Hammond, 2015b; Theriault, 2015). At nearly $200,000, Saxton's salary was close to two and a half times what his predecessor made.[20] During the 2015 legislative session, with the encouragement of Governor Brown, OEIB was abolished, and in August, Brown announced the retirement of CEO Nancy Golden. It remains to be seen whether the position of CEO will continue without Golden and minus OEIB (Jaquiss, 2015; Hoffman, 2015).

Recent policy history: Local flexibility and a legislated focus on college completion. Oregon's recent education policy history illustrates the state's encouragement of local control in the development of approaches to address state policy. A decade before No Child Left Behind (NCLB), Oregon began efforts to develop state standards and aligned assessments, putting the state ahead of others. Oregon has as history of allowing districts flexibility in the choice of assessment methods, permitting a combination of assessments including work samples and portfolios (Louis, et al., 2009). Recently, along with the adoption of the Common Core State Standards (CCSS), Oregon has moved from the Oregon Assessment of Knowledge and Skills (OAKS) to Smarter Balanced Summative Assessments, a testing system from one of two national assessment consortiums that provide CCSS-aligned exams. Participants in this study expressed frustration over the switch to new tests and to their use in teacher evaluation. Regarding

[20] Noor took office July 1, 2015. I was not able to obtain information about his salary.

teacher quality, Oregon has long held a commitment to supporting new teachers, and a preference for local control is evident in the state's teacher and administrator mentor program, initiated in 1993 as part of the "Oregon Educational Act for the 21st Century" (Louis, et al., 2009). The mentor program provides competitive grant funding to districts to develop local, research-based mentoring and induction programs, with few specifications regarding program design. Having undergone slight revisions since its inception, the program remains in place and continues to be funded (Oregon Department of Education, 2015e).

The current centerpiece of Oregon's efforts to improve the education system is another Kitzhaber-endorsed 2011 law, which established Oregon's "40-40-20" goal. This law pressures the education system to ensure that by 2025, 40% of the state's adult citizens will hold a college degree, 40% will hold an associate's degree, and 20% or fewer will have earned only a high school diploma. Part of the initial work of OEIB was the establishment of milestones for student success in relation to the 40-40-20 goal, including kindergarten readiness, early literacy targets, improvements in student attendance, progress toward graduation for ninth graders, and improved graduation rates (Oregon Secretary of State, 2015). Among ODE's strategic priorities related to the 40-40-20 goal are the successful implementation of the new teacher evaluation policy and a commitment to link evaluation results to professional development supports (Oregon Department of Education, 2015d).

Key stakeholder organizations: Collaboration marks policy work. Reflecting the state's participatory political culture, education policymaking in Oregon has become a

collaborative affair over the past twenty years, with broad involvement from education groups as well as business representatives and others. This study confirms that near circle players like OEA and COSA collaborate regularly, not only to lobby at the legislature, but also to discuss and develop policy (Louis, et al., 2009). With 45,000 members, OEA, an affiliate of the National Education Association (NEA), represents nearly all Oregon teachers, while American Federation of Teachers Oregon represents teachers in only four districts (Oregon Education Association, 2015; American Federation of Teachers, 2015). COSA, as its name suggests, is an umbrella organization. Four school administrator organizations are considered departments of COSA: Oregon Association of School Executives (OASE), Oregon Association of Secondary School Administrators (OASSA), Oregon Elementary School Principals Association (OESPA) and Oregon Association of Central Office Administrators (OACOA). In addition, COSA is affiliated with several other organizations that represent a range of groups, from Latino/a administrators to student councils.

A fairly new player to the education policy landscape in Oregon is the Chalkboard Project, which was founded in 2004 by a group of Oregon's leading philanthropic foundations. Chalkboard, which describes itself as an "education transformation organization," provides independent research, partners with educators and schools, and advocates for favored policies (Chalkboard Project, 2015e). The organization began its work in 2004 by interviewing educators across the state to learn about how the education system was working and what changes might improve things (Louis, et al., 2009). Chalkboard prides itself on bringing voice to the experiences and concerns of practicing

educators (Chalkboard Project, 2015d). The organization is affiliated with the Policy Innovators in Education (PIE) Network, a national reform advocacy group aimed at raising academic standards, closing achievement gaps, increasing educator effectiveness, advancing school choice, and increasing school accountability. Other groups affiliated with PIE include the Center for American Progress, the National Council on Teacher Quality, Students First, and the National Alliance for Public Charter Schools (Policy Innovators in Education Network, 2015b). The PIE Network is considered to be a leader among education reform advocacy organizations (ERAOs), which generally include teacher accountability policies among their favored reforms (McGuinn, 2012a).

Chalkboard's primary initiative is the Creative Leadership Achieves Student Success (CLASS) Project, which the organization began in 2007 by funding three Oregon school districts to reform educator evaluation, improve professional development, and expand compensation models and career paths for teachers. Districts received start-up funding along with an amount of funding per teacher for implementation, totaling up to $250,000 a year (Dungca, 2012). Chalkboard staff work closely with partner districts, assisting in the development of locally designed initiatives to meet the expectations of the program. Over the years, the program has grown to include over 40 districts, and Chalkboard boasts, "Nearly 40% of all Oregon teachers and students are in a CLASS district" (Chalkboard Project, 2015a).

Education outcomes: Low state rankings cause concern. Governor Kitzhaber's efforts to reform Oregon's education system came as no surprise in a state that consistently grapples with education outcomes, particularly high school graduation rates,

where Oregon ranks worst among U.S. states with only 69% of high school students graduating in four years (Editorial Projects in Education, 2015a). (See Appendix H for data on Missouri and Oregon education characteristics, outcomes, and ratings.) Participants in this study noted the state's low performance, indicating that the teacher evaluation policy is one attempt to improve student success. On some indicators, Oregon performs about average, including on the National Assessment of Educational Progress (NAEP), where the state performs similarly to the national average (National Center for Education Statistics, 2015). However, per pupil education expenditures are more than $1,000 below the national figure, 33rd among the states. Moreover, Oregon received a grade of C- and a state ranking of 41st in Education Week's latest *Quality Counts* report (Editorial Projects in Education, 2015b). The report's rankings take into consideration a variety of indicators under three broad umbrellas: K-12 achievement, school finance, and chance for success (parent and family factors, early education, and indicators of preschool through adult participation and success). In addition to graduation rates, Oregon ranked low among the states on preschool enrollment, steady employment among adults, and improvements in NAEP scores, particularly on reducing the gap in scores between poor and non-poor students. On the other hand, eighth grade achievement on both NAEP reading and math exams compared well against other states.

In comparison to Missouri, Oregon has more nonwhite students, more English learners, and more students eligible for free or reduced-price lunch (U.S. Department of Education, 2015a). Oregon was ranked significantly lower than Missouri (35th compared to Missouri's 23rd) on education in the 2015 *Kids Count* report from the Annie E. Casey

Foundation. An attempt to measure child-well being in U.S. states, the report's education rankings represent four key predictors of success: preschool attendance, fourth grade reading proficiency, eighth grade math proficiency, and on-time high school graduation (Annie E. Casey Foundation, 2015).

History of Teacher Evaluation Policy and Practice in Oregon

Teacher evaluation policy: A policy in place for over 35 years. In 1979, before *A Nation at Risk* and other reports criticizing American schools and encouraging improvements in the evaluation of teachers, the Oregon state legislature passed the state's first teacher evaluation law. The law required the annual evaluation of both probationary and post-probationary teachers, termed "contract teachers" in Oregon (Or. Rev. Stat. § 342.805). Districts were required to adopt performance criteria; to conduct multiple observations, each of which was to be preceded and followed by a conference between the teacher and the principal; and to implement an improvement program for those teachers who needed support. In addition, the law mandated that local teacher evaluation systems be developed "in consultation with" administrators and teachers, to include some teachers appointed by the exclusive representative of the teachers (the teachers association or union) (Hungerford & Dickson, 2012). This policy language requiring administrators and teachers to work together on local design reflects the tendency of moralistic cultures to encourage broad participation in policy efforts. The teacher evaluation law was amended in 1997, removing the requirement that evaluations of contract teachers be conducted annually and adding the opportunity of "peer assistance" for teachers on improvement plans (Hungerford & Dickson, 2012; Reed, 1988; Oregon

Revised Statutes, 2015). There were no further changes to the law until Senate Bill 290 (SB 290) was passed in 2011, ushering in Oregon's new teacher evaluation policy.

Teacher evaluation practice: Nominal compliance with no state authority. Prior to the adoption of the current teacher evaluation policy, there was a lack of uniformity in teacher evaluation practice across Oregon and little guidance from the state. Nominal compliance resulted in some teachers going years, even decades, without being formally observed. Several Phase I participants commented on the status of teacher evaluation practice prior to the development and implementation of the new policy, and all comments indicated that the new policy is an improvement over past practice. One district administrator stated that the new system is "ten times better than what we had before." He went on to explain his view that the new approach brings a necessary focus to effective classroom practice:

> Our evaluation systems – many of them were just archaic and they just they really didn't focus on a teachers' ability to help educate kids. They focused simply on a teacher delivering instruction in a way and not worrying about whether the kids got it or not.

Others commented that the state lacked a model teacher evaluation system, and they explained that there was also a lack of support and guidance from the state, along with a lack of accountability. According to another district administrator, this lack of accountability made it common for teacher evaluation to get lost among competing demands for principals' time:

> I certainly think everybody being evaluated on a regular basis is really nice. And that's not something that's always happened, not only in our district, but in the state. There hasn't been a real strict measurement of that or accountability for that . . . And I think it's pretty fair to say across the state because I've talked to a lot of H.R. [human resources] people who say, oh my gosh I'm so frustrated – you get to the end of the year and the principal says, oh gosh, I didn't get any of my evaluations done. They really can't do that anymore.

State-level participants echoed the perspective that teacher evaluation practice in Oregon was weak prior to the development of the new policy. A telephone survey of school district human resource directors was administered by ODE in 2010 to learn about the status of teacher evaluation across the state. The sample included 63 of the state's 197 school districts, representing size and geographic diversity (Oregon Department of Education, 2012). According to the representative of ODE, the survey illustrated that there was a "vast, vast difference" in how districts approached teacher evaluation, with no consistency in terms of evaluation criteria, frequency, or measurement. The representative of Chalkboard expressed the belief that there was a lack of quality in the teacher evaluation practices of most districts, even some of the larger, more affluent districts, noting that many were still using "the real old checklist" systems. As explained in the review of the literature, such systems began to appear in the 1950s and have undergone criticism since the 1980s (Marzano, et al., 2011; Darling-Hammond & Berry, 1986; Hallinger, Heck &Murphy, 2014; Peterson & Kauchak, 1982).

State-level actors echoed the perspective of local actors that there was a lack of direction from the state, which resulted in a lack of consistency in teacher evaluation practice, and some teachers went many years between formal observations of their classroom practice. The representative of the teachers association explained:

> I've met – in my travels around the state – have met people who had, prior to Senate Bill 290, not been evaluated their entire career outside of their probationary years. We are talking 20 years later, 30 years later. Or who had been evaluated but really it was more of a, you know – I haven't even been in your classroom but I'm signing the documentation saying you've been evaluated; sign this form so you can keep your job – you know, kind of a contract extension type interaction as opposed to an actual evaluation where there's a professional growth opportunity.

This participant went on to explain that an over emphasis on formal observations led to what is often referred to as teachers presenting a “dog and pony show” when the principal is present, a portrayal that the participant feels is both an accurate and inaccurate depiction of teacher evaluation practice in Oregon in the years prior to the development of the new policy:

> It's accurate in the sense that, of course if you know when your supervisor is going to be watching your work, you’re going to put on your best, are going to put your best foot forward, which is a good thing – that's not a bad thing. Of course you would do that. The funny part about that is that there have been, you know, kind of street – stories from the streets – that are things like the myth of the

> teacher doing the same lesson every year . . . I don't really think that everyone is such a terrible teacher that the dog and pony show is the only version of good teaching that they display and then they go back to some sort of horrifying practices, but I do think that more aptly there's no mechanism in that type of system for people to actually create an instructional leadership relationship with their supervisor where they are trusting that person to give good feedback.

Development of the current policy: Coordinated state efforts meet federal hurdles. From the inception of the policy through the process of policy interpretation and implementation planning, state-level stakeholder groups worked together. Their work was not without difficulty, however, especially as regards addressing the expectations of the ESEA Waiver program. Oregon educators began meeting to discuss potential changes to the state's teacher evaluation policy in 2010. A work group was convened by OEA, which included teachers who held OEA leadership roles and OEA staff, along with school administrators, college and university faculty, and other stakeholders. The group concluded that teacher evaluation should be part of a system of support across the career continuum from pre-service preparation, through the mentoring and induction period, and across the career. The group expressed the value that teacher evaluation should promote professional growth and be supported by positive teaching and learning conditions (Oregon Education Association, 2012).

According to state-level actors, each of their organizations – ODE, OEA, COSA, and the Chalkboard Project – supported the 2011 bill that would become the centerpiece of Oregon's new teacher evaluation policy, SB 290, and worked toward its passage.

When asked who was responsible for moving the policy agenda forward, participants credited state senators from both parties as well as the Chalkboard Project. In response to this question, Chalkboard's representative explained that the work on the bill involved collaboration across groups and that these same organizations continued to work together throughout the policy's implementation:

> I think that any organization you talk to would say it was them because I think – and I could be biased because I'm here [at Chalkboard] – our story is that Chalkboard did it, but we knew that we had to do it in a coalition . . . So I think the fair thing would be, like who really cares who started this conversation, everybody finished it. So I think that everybody would be Chalkboard, the Oregon Education Association, and to a little bit lesser but an important role would've been COSA, our Coalition of Oregon School Administrators . . . We were the ones who stayed at the table for supporting the department with implementation.

With the exception of the Chalkboard Project, which has only existed since 2004, the stakeholder groups involved in the development of Oregon's teacher evaluation policy are all near or inner circle players as identified by Louis, et al. (2009) in their study of Oregon's state education policy culture.

Reflecting the influence of the broader policy environment, the ODE representative acknowledged that part of the reason for the development of a new teacher evaluation policy in Oregon was the national conversation around educator effectiveness that was prompting U.S. states to reconsider existing policies on educator evaluation:

> I think being on the national agenda, it was talked about across the country. I think that probably enhanced the work. It elevated awareness of evaluation systems. It wasn't just something that Oregon was doing; it really was a national movement. I think that helped move it forward. Even the support of the Oregon Education Association – they were listening to, you know, their national counterparts and so everybody was aware that this was moving forward. And without the federal policy and if it was just Oregon, I think it wouldn't probably have.

As mentioned above, OEA had convened stakeholders for discussions about teacher evaluation in 2010, and the OEA representative confirmed that the state teachers association was influenced in their approach to teacher evaluation by their national affiliate, the National Education Association (NEA), which had issued a series of statements addressing teacher evaluation beginning in 2007. In their publications, NEA emphasized the need for standards-based teacher evaluation systems that focused on professional growth and included multiple measures, job-embedded professional learning, and strategies to promote student learning and development (National Education Association, 2015b). According to the OEA representative, another reason for the association's participation in policy development was the fact that early drafts of SB 290 included "anti-labor practices" related to seniority and contractual rights. These elements were removed before SB 290 was introduced at the legislature. According to a document released in 2012 by the Chalkboard Project, there had existed in Oregon for decades a perception that teacher contract rights were so strong in most school districts that it was

difficult if not impossible to remove a teacher from her position for poor performance (Hungerford & Dickson, 2012).

SB 290 was passed by the Oregon state legislature in 2011, a bi-partisan bill that augmented the 1979 teacher evaluation law; it outlined new requirements for the evaluation of both teachers and school administrators. The bill passed overwhelmingly, earning praise from both Democrats and Republicans in the legislature as well as then Governor Kitzhaber, a Democrat (Melton, 2011b). SB 290 directed the Oregon State Board of Education to adopt teacher and administrator standards that were to include multiple measures of effectiveness and evidence of student academic growth, and that would strengthen educator's practice, provide opportunities for professional growth, and allow for individual goal setting based on the needs of the educator and his context. In addition, the law required that in each school district, administrators, teachers, and the exclusive bargaining representative of the teachers would work collaboratively to develop an evaluation system that would meet the newly established standards (Oregon Department of Education, 2014a; The Oregonian, 2015). The move from the term "consultation" in the 1979 legislation to the term "collaboration" in the 2011 bill was interpreted by Oregon's Office of the Legislative Counsel to require a higher level of interaction among the parties, though it stopped short of requiring mutual agreement between the school district and the exclusive representative of the teachers (Hungerford & Dickson, 2012). The SBE adopted the teaching standards of the Interstate Teacher Assessment and Support Consortium (InTASC) as Oregon's model teaching standards in December 2011, amending the state's administrative rules on personnel policies for

public elementary and secondary schools (Or. Admin. R. 581-022-1720). (For a summary of Missouri and Oregon's teacher evaluation policies and systems, see Appendix I.)

ESEA Flexibility Waiver: Back-and-forth with USDOE. By September 2011, one month after U.S. Secretary of Education Arne Duncan announced plans to offer flexibility to states under NCLB through the Elementary and Secondary Education Act (ESEA) Flexibility Waiver Program, Oregon had begun plans to develop a waiver application (Oregon Department of Education, 2011). ODE assembled a work group to address the principles of the waiver, including Principle Three on educator evaluation. The ESEA Waiver Educator Effectiveness Workgroup included 24 individuals representing ODE, OEA, COSA, the Chalkboard Project, schools and districts, colleges and universities, and additional state-level stakeholders. About two-thirds of the group's members were practicing K-12 educators, both administrators and teachers (Oregon Department of Education, 2014a). This group, which began meeting in fall 2011 to develop the waiver application and continued its work throughout the 2011-2012 academic year, grew in membership to about 50 as the group developed the Oregon Framework for Teacher and Administrator Evaluation and Support Systems (Oregon Framework). In addition to involving large work groups in the design of the waiver application, the state sought and received feedback from hundreds of educators about the plan (Castillo, 2012).

When the work on the waiver application began, a crosswalk was conducted that demonstrated close alignment between Oregon's teacher evaluation law and the requirements of the waiver. State-level stakeholders agreed that the teacher evaluation

requirements outlined in SB 290 aligned quite nicely with the requirements of the waiver, with one notable exception. Although Oregon's legislation required that teacher evaluation systems include evidence of student learning, the language was not as stringent as the federal requirement that evaluation systems include data on student growth as a "significant factor" (U.S. Department of Education, 2014). This would prove to be a difficult sticking point in negotiations between ODE and the U.S. Department of Education (USDOE) as the waiver application was developed, approved, and renewed over the coming months and years.

When asked about Oregon's experience with the waiver process, the representative of ODE said with a laugh, "That's been an interesting journey!" According to state-level participants, early on, tensions arose between Oregon stakeholders and USDOE around the student growth requirement, with USDOE repeatedly expressing dissatisfaction with Oregon's approach to addressing the requirement. The design included in Oregon's initial waiver application, submitted in early 2012, allowed local school districts to determine their own approach to meeting the requirement. USDOE did not approve Oregon's original waiver request, citing concerns about the state's ability to meet established timelines and to ensure the validity of measures within the evaluation system, particularly measures of student growth (Yudin, 2012). ODE submitted a revision to its educator evaluation plan in October 2012, bolstering its requirements for student growth measures in local plans. Oregon received conditional approval of its waiver application in July of that year, and it was expected that the state would pilot various approaches to meeting the student growth requirement and finalize its evaluation system

by the end of the 2012-2013 school year (Duncan, 2012). Oregon's preference for local control in the design of the student growth requirement reflects a tradition of allowing schools discretion in executing the particulars of education policies (Louis, et al., 2009).

After a year of piloting student growth models, in August 2013 Oregon requested another year to finalize its teacher evaluation system. As a result, the state was placed on high risk status by USDOE and given one more year to address the federal government's concerns. With the high risk designation came intensive monitoring from USDOE (McNell, 2013; Delisle, 2013). According to state actors, the placement on high-risk status was mainly due to the state's failure to meet timelines, and not due to the quality of work underway. Participants expressed frustration with USDOE's approach. The OEA representative stated:

> We were taking things really slowly, and very purposefully, and trying to be very meaningful in our path toward implementation, and the feds didn't like that a lot. The U.S. Department of Ed. was very much – you need to get this done, you needed to get it done by this date, you haven't got it done, you are now on high risk – very the opposite of good action research and the opposite of any type of implementation science you would look at that says, hey, you should take this slow.

It seems that the participatory, collaborative approach favored in Oregon – a cornerstone of how education policy gets done in a moralistic state – led to a disappointing result when it met up against federal pressures.

During the time when they were on high risk status, Oregon leaders worked closely with USDOE personnel to arrive at a solution for ensuring that student growth represented a "significant" portion of a teacher's evaluation rating: this solution was the Oregon Matrix. This matrix is a model by which an evaluator combines scores from the professional practice measures in the teacher evaluation system with scores from measures of professional responsibilities, and then places those scores into a matrix, in relationship with the ratings on student growth measures, to determine a final rating for the teacher. According to the representative of ODE, the Oregon Matrix ensures that student growth measures account for at least 20% of a teacher's rating, something that Oregon was asked to guarantee in their work with USDOE. Again, despite the fact that they believed they were meeting expectations expressed by USDOE in their work together, Oregon received only conditional approval in its request for an extension of its waiver. The representative of ODE took on a discouraged tone in describing the situation:

> They [USDOE] were with us all the way through and knew what we were doing and said, yeah, that we were on the right track. Then in the very end when we submitted our final model, our matrix in which they, you know pretty much agreed to, they came back again and said we don't quite think you have created enough rigor and consistency around this state assessment piece.

At the time of Phase II interviews for this study, Oregon was still under a conditional waiver, having submitted its application for waiver renewal in March 2015. State-level stakeholders had gone back to the table with USDOE to agree on a new approach to measuring student growth and calculating its impact on a teacher's

summative rating. The centerpiece of the approach outlined in the March 2015 waiver renewal submission was "student growth percentiles." This approach, applicable only to teachers whose students take mandated state tests, measures growth for each student by comparing a student's change in academic achievement with that of his or her "academic peers." Academic peers are defined as Oregon students of the same grade level with a similar history of results on state achievement tests (Oregon Department of Education, 2015c). A teacher's student growth score is then calculated from his students' student growth percentiles.

State-level participants expressed concerns about the new approach, both the validity of the approach itself and how it was to be received by local educators; they anticipated complaints about yet another change to the system and objections to an approach that places so much emphasis on state mandated achievement tests. One participant stated, "We just resubmitted our waiver, and that has a brand new component that no one even knows about yet. To be blunt, you know, the proverbial [expletive] is about to hit the fan on that one." USDOE finally accepted Oregon's approach to the student growth requirement when it granted full, unconditional approval of the state's waiver renewal request in July 2015, more than three years after the state submitted its initial waiver application (Whalen, A., 2015).

The current policy: A "tight-loose" model. Oregon's teacher evaluation policy combines the requirements of SB 290, which were an amendment to the 1979 teacher evaluation law, with the requirements of the ESEA Flexibility Waiver. The policy is expressed in the Oregon Framework for Teacher and Administrator Evaluation and

Support Systems (Oregon Framework), developed in 2011-12 and updated regularly.[21] The framework outlines requirements for local teacher evaluation systems using a "tight-loose" model, whereby the state establishes tight guidelines for expected outcomes, but a loose approach in letting district determine the strategies to achieve those goals (U.S. Department of Education, 2012). The framework details the following required elements for local teacher evaluation systems: standards of professional practice, differentiated performance levels, multiple measures, an evaluation and professional growth cycle, and aligned professional learning. In addition, the framework stipulates that districts are required to develop their systems in collaboration with administrators, teachers, and their exclusive bargaining representatives, and that districts must use the evaluation system for personnel decision-making (Oregon Department of Education, 2014a). Districts were required to submit plans for review in July 2013.

As the policy was being rolled out, ODE and its stakeholder partners reviewed various performance rubrics for assessing teachers on the model teaching standards, conducting crosswalks for each. Four performance rubrics were recommended, and districts that chose one of those rubrics were deemed to have satisfied the state's guidelines.[22] If a district utilized a different performance rubric or adapted one of the recommended rubrics, it was required to submit a crosswalk to ODE for review and approval (Oregon Department of Education, 2015b). Beyond the specificity regarding teacher performance rubrics and the above mentioned precision with which ODE has

---

[21] Oregon's policy encompasses both teacher and administrator evaluation, but in this paper discussion is limited to teacher evaluation.

[22] The four recommended rubrics are: Framework for Teaching from the Danielson Group, Kim Marshall Teacher Evaluation Rubrics, Marzano Teacher Evaluation Model, and the Salem-Keizer School District Licensed Staff Assessment and Evaluation Rubric (Oregon Department of Education, 2015b).

been required by USDOE to spell out requirements for meeting the student growth mandate, there was considerable flexibility afforded to school districts in determining a local teacher evaluation approach.

Teacher evaluation funding causes concerns over equity. An uneven distribution of funding for teacher evaluation caused concerns in Oregon, and state-level stakeholders were especially frustrated by this reality. One state-level participant described teacher evaluation funding in Oregon's as a "bipolar landscape," noting serious concerns about a situation of haves and have-nots. Ultimately, three sources of funding supported school districts – the legislature, grants from CLASS and/or ODE, and federal Teacher Incentive Funds (TIF).

Legislative funding. The Oregon state legislature did not provide any funding when it passed SB 290 in 2011, leaving Oregon's 197 school districts with no new monies to implement the law's requirements. However, districts that sent teams to the regional training sessions held during 2012-2013 received reimbursement from ODE for expenses, including the cost of hiring substitute teachers. During the 2012 legislative session, HB 3233 was passed, funding the Network for Quality Teaching and Learning (NQTL), which is led by ODE and OEIB. The legislation allotted funding to promote a variety of initiatives related to effective teaching and leadership, including teacher evaluation (Oregon School Boards Association & Coalition of Oregon School Administrators, 2013). Monies that were awarded for teacher evaluation were part of a category of funds intended to assist districts in implementing both educator evaluation systems and the CCSS. At the time of this writing, 179 of Oregon's 197 school districts

had received NQTL funding. Awards ranged from less than $1,500 for the smallest of these districts (with an enrollment of fewer than 20 students) to over $500,000 for the largest of these districts (with an enrollment of over 45,000 students) (Oregon Education Investment Board, 2015b). In 2013, NQTL funds along with monies from another legislative initiative, the Strategic Initiatives for Student Success, were distributed to small or rural districts in recognition of their lack of capacity for implementing major policy initiatives underway at the time, including teacher evaluation. According to information at ODE's website, funds were distributed to 129 small districts with enrollments ranging from about 150 to 2,000. Each of these districts received between $2,500 and $11,000 to support their work on educator evaluation and other initiatives (Oregon Education Investment Board, 2015b). Finally, during the 2013 legislative session, $6 million from NQTL funds was designated by the legislature to fund educator evaluation activities in *all school districts* for the 2013-2014 and 2014-2015 academic years (personal correspondence, June 20, 2015).

CLASS and Collaboration Grants. As described earlier, Chalkboard funded districts through its CLASS project to support work related to educator evaluation (and three other areas) beginning in 2007. In 2011, the state legislature passed Senate Bill 252, establishing the School District Collaboration Grant Program (Collaboration Grants), to be administered by ODE. The requirements of this grant program are essentially the same as Chalkboard's CLASS program, and now, instead of providing funds directly to districts through CLASS, Chalkboard was contracted to provide technical assistance to the Collaboration Grant districts. The legislative appropriation for the Collaboration

Grants was $5 million over the 2011-2013 biennium (Oregon School Boards Association & Coalition of Oregon School Administrators, 2011). The legislation also provided that the Collaboration Grant districts would pilot the state educator evaluation guidelines (Oregon Department of Education, 2012). Through a competitive application process, 12 districts were awarded Collaboration Grants in 2012-2013, and these districts piloted the Oregon Framework, along with two additional pilot districts that were funded by OEA. Six of these 12 districts were awarded Collaboration Grant funds again in 2013-2014, along with nine additional districts and two regional consortia, one that included three districts, and one that included eight districts. These consortia were sponsored by two of Oregon's Education Service Districts (ESDs). Those districts awarded Collaboration Grants in 2013-14 were responsible for piloting student learning and growth goals (personal correspondence, June 20, 2015). ODE has continued to award Collaboration Grants, with funds going to support multiple efforts, including teacher evaluation.

Teacher Incentive Fund. Prior to the passage of SB 290 and a few years after initiating CLASS, Chalkboard was awarded federal TIF funds in 2010 for work in six of the already established CLASS districts, receiving $24.4 million over five years. Ultimately, two of these six districts dropped out of the TIF initiative, forgoing the federal financial support. The TIF districts were chosen based primarily on their status as underperforming schools with a large percentage of high-need students, requirements of the program. Districts involved in Chalkboard's TIF grant were obligated to shape their teacher evaluation and compensation systems to recruit, retain, and reward teachers deemed to be effective (Chalkboard Project, 2015a). One state-level participant

emphasized the size of the TIF grants, stating, "The real money [for teacher evaluation] was federal money that came in through TIF."

Perspectives of state-level policy actors about funding for teacher evaluation. Though all of the state-level participants expressed concerns about funding for teacher evaluation efforts, it was the representative of COSA who registered the most serious concerns. For this participant, issues of funding equity were especially troubling. Regarding the monies districts received for teacher evaluation efforts, the COSA representative explained that about 15 percent of districts had money through Chalkboard and/or TIF to begin teacher evaluation work preceding and directly following the passage of SB 290, and in these districts, "People could dream and create with no restrictions." However, the participant explained that in the remaining 85% of districts, there was no funding at the start of the effort, limiting the ability of most districts to take the time to develop a local plan thoughtfully. The COSA representative expressed skepticism about the funding associated with Chalkboard's efforts, characterizing the organization as being motivated by accountability and reform. As stated earlier, 40 percent of Oregon teachers and students are in CLASS districts. This includes those districts that initially received Chalkboard funding through the CLASS project, those that received federal TIF funds, and districts that received state-funded Collaboration Grants (Chalkboard Project, 2015a). The Chalkboard representative affirmed that districts that received grant funding were able to design and implement the policy "as intended" and that "Not every district had that ability."[23]

---

[23] Information about the amount of funding that Chalkboard itself has provided directly to districts through CLASS was not made available by the organization.

Another concern had to do with geography. Almost all of the districts that have received substantial, direct funding through Chalkboard, TIF, or the Collaboration Grants are in the Western half of the state.[24] The representative of OEA indicated that districts in the Western part of the state were advantaged as the policy was being rolled out – even if they had not received grant funding – because of their proximity to CLASS districts, stating, "There were a lot more things to look at if you had a neighboring district that you trusted and knew, if you were in the valley side of Oregon." Certainly the situation has improved over the past couple of years as the NQTL funds meant that all districts had access to at least some funding for teacher evaluation activities, but the fact remains that early on, a small percentage of districts had new monies to commit to the effort, and for those districts receiving TIF funds, the injection of funding was substantial.

Three Oregon School District Cases

As explained in Chapter Three, three small Oregon school districts comprised of at least three school sites were chosen for this study, based on their approach to teacher evaluation system design. The three districts were: Meyers Grove School District, a district that adopted a state-approved teacher evaluation model; Davies School District, a district that developed its own teacher evaluation model, and Nilsenville School District, a district that worked with a state-level direct support organization to develop a teacher evaluation model.[25] (For information on the characteristics of Phase I participants, see Table 2, Chapter Three, p. 91.)

---

[24] The exceptions are small rural districts that are part of consortia organized through the ESDs, and these districts have been added to Chalkboard's work in the past three years (Oregon Education Investment Board, 2015b).

[25] Names of participating school districts are pseudonyms.

Meyers Grove School District. Through a process characterized by collaboration and high trust, educators in Meyers Grove School District chose to adopt one of Oregon's state-approved teacher evaluation rubrics. A fairly small teacher evaluation committee worked well together, committed to evaluating and modifying the plan as implementation moved forward.

The community and its schools: A rural, low-income district. Set in Northwestern Oregon, Meyers Grove School District covers a large area, incorporating multiple small communities. Schools in the district are spread out; with at least half an hour's drive between the district's furthest flung schools. The recreation and tourism industries are major employers here, and the median household income is considerably below the state figure of $50,229. Per pupil expenditures in 2013-2014 were significantly higher than the state average of $10,256. (For data on community characteristics, district characteristics, district outcomes, and district educators for all three Oregon districts included in this study, see Appendix K).

At both the elementary and high school levels, Meyers Grove met Annual Measureable Objective (AMO)[26] targets in reading but did not meet these targets in mathematics. At the middle level, the district met targets in both areas. In 2014, the percentage of Meyers Grove students who met or exceeded state standards on OAKS testing was above the state average on about half of the ten exams. The 2014 four-year graduation rate in Meyers Grove was high – more than 15 points higher than the state rate

[26] As mentioned in the previous chapter, Annual Measurable Objective is defined as the goal set by each state indicating the minimum percentage of students who must meet or exceed standards as measured by the state's achievement exams (U.S. Department of Education, 2015a).

of 69% – but only about half of Meyers Grove graduates go on to post-secondary education.

In 2012-2013, the average salary of teachers in Meyers Grove was very near the state average, considerably higher than the average teacher salaries in the two other Oregon districts in this study. Most of the school and district administration has been in place for over five years, with the exception of the superintendent who came to the district within the last few years.

Teacher evaluation decision-making: Respected leaders inspire teamwork. In Meyers Grove the teacher evaluation committee was smaller than most others in this study, but it did include representation from district administration, building principals, teachers, and the school board. The density of policy collaboration was high in this district (see definitions of terms in Chapter One), with the committee meeting at least monthly beginning in the fall of 2012, engaging directly with state-level stakeholders, and seeking feedback from district staff regularly. In addition, administrators and teachers praised the cooperative nature of the process – the committee was truly a team. Although it was reported that teachers in Meyers Grove, like teachers in most other districts in this study, were fearful about the adoption of a new approach to teacher evaluation, it seemed that this fear was limited to a small number of teachers; team members were committed to helping their colleagues understand the teacher evaluation plan and alleviating their fears.

Participants made it clear that much of the credit for the high degree of collaboration and trust that accompanied the decision-making process goes to the

superintendent. Participants considered him to be resourceful, collaborative, not a "know-it-all." To the superintendent himself it was very important to take an active role in the development of the local teacher evaluation approach; he felt that it was his responsibility to ensure that teachers and administrators developed the system collaboratively together. According to one of the teacher interviewees, all viewpoints were respected:

> In our committee, it was extremely collaborative. None of us felt that somebody was in charge telling us what to do . . . And we did not feel that we were being run rough shod by the administration at all.

Another factor in the success of the local process was the feeling of trust between the district administration and the teachers union. It was noted that the teachers on the committee were strong union members, and the administrators who were interviewed credited those teachers with keeping the process positive, constructive, and non-confrontational. As one administrator said, "Working with the union reps was a very productive piece of it. And we feel that the collaborative work that we did has really paid off."

There was a sense that the teacher evaluation committee dug deep and sought to understand the policy beyond the mechanical level. Likely this was due to the district's commitment to involve *all* members of the committee in attending training sessions at both the regional and state level; information received by the committee was not filtered through administration, all members had opportunities to interact with state-level stakeholders, and all were committed to sharing resources with one another. Compared to

other districts in this study, Meyers Grove participants exhibited a more thorough understanding of the state policy and its requirements, timelines, and nuances.

Participants indicated that the two primary external sources on which the committee relied for information were ODE and Oregon school districts that had piloted the state-approved rubrics. Also, members of the Meyers Grove teacher evaluation committee expressed appreciation for the state-level collaboration among ODE, OEA and Chalkboard on joint training sessions for school districts. They felt that the interests of all parties were protected in this way. As a Meyers Grove principal stated:

> There were those three entities – the department of education, and then the teachers union, and the Chalkboard Project . . . It always gave you that balanced perspective then; you didn't just get a teachers' perspective or an administrators' perspective.

After conducting a close review of Oregon's state-approved rubrics, Meyers Grove chose to use the Kim Marshall teacher evaluation rubrics, which had been effectively piloted in Oregon and approved by the state. In choosing among the approved state rubrics, there was an emphasis on finding the right "fit" for the district's culture. The district conducted a local pilot during 2013-2014, and the plan was fully implemented in the fall of 2014. After the initial pilot year and again during the first months of full implementation, the committee was purposeful in evaluating the model and making modifications to the plan. During the period when interviews were conducted, committee members were preparing a self-assessment of their evaluation system for Oregon's

teacher evaluation peer review process, and they expressed the intention to go back and tweak their system as result of this exercise.

Despite Meyers Grove's praise for the harmonious process of local decision-making, participants did express two key concerns. The first concern, repeated by several participants, had to do with the district's size; committee members felt that, because the district is very small, the development and implementation of a system for meeting the requirements of a new, broad-ranging policy was especially burdensome. In addition, a concern was expressed about the intense time commitment that was required as a result of the state's local control approach. In Oregon, even districts that adopted a state-approved rubric faced a considerable amount of design work in developing the remaining components of the system. One participant stated: "Ideally the state would have said, 'Okay, here. This is the document you are going to use.' . . . There's a lot of time and energy involved that, if it was streamlined, it could have been more efficient." However, even this participant praised the committee's work and felt that their efforts resulted in a good local system.

Perspectives on the policy: Keyed into policy context. Participants expressed frustration about shifting state policy requirements, though they seemed to understand that this was a result of issues regarding the renewal of Oregon's federal ESEA waiver. Particularly frustrating was the move to Oregon's matrix model for calculating summative evaluation ratings. Also, there were concerns about ODE's efforts, with participants noting delays in information being shared with districts and getting posted on the ODE website. The principal interviewee put it this way: "The roll out was so fast that

it was hard for them, I think, to stay up on top of materials for us to access . . . So I think they're kind of creating it as they go."

In describing the policy's goals and the reasons for its adoption, district administrators, principals, and teachers alike expressed a clear understanding of the link between Oregon's teacher evaluation law and the policies of the Obama administration. Participants viewed the development of a standardized system of teacher evaluation as a central goal of the policy, and teacher evaluation documents provided by the district indicate the design team's emphasis on linking evaluation to professional development. In addition they saw the policy to be aimed at improved teacher quality.[27] Participants did not view these policy goals as conflicting. However, they did identify conflict between the political discussions about the policy and what the policy means to educators on the front lines. Participants expressed frustration that some members of Oregon's state legislature link teacher evaluation to a business mindset about schools and educators. A principal stated:

> It's really hard when politicians try to tell educators how to do their job. Sometimes what works in business and industry as a nice, clean-cut model is not working, and you can't control the product coming in to you. I don't think you can always take a business model and apply it to education.

Also, educators in Meyers Grove were frustrated by the media's tendency to focus on the use of student test scores in reporting about teacher evaluation. One teacher articulated this view: "In the media I have heard different things – more of an emphasis on holding

---

[27] As defined in Chapter One, the term teacher quality has to do with measurable teacher attributes and bears an accountability focus.

teachers accountable to certain standardized test scores, and that being the emphasis rather than truly improving education." Despite these concerns, committee members reported being unified around the policy goal of professional growth throughout their process of selecting, piloting, and implementing the teacher evaluation plan.

When queried about their views on the value of the state policy and the strengths of their local approach, participants emphasized the standardization of teacher evaluation within Meyers Grove and across the state of Oregon, which they viewed as ensuring both consistency of evaluation practice and accountability for school administrators as evaluators. In articulating the challenges of developing a local approach to teacher evaluation and concerns about their own policy moving forward, participants displayed a fairly sophisticated understanding of potential implementation difficulties. They expressed concerns about the fidelity of implementation, especially in terms of applying the policy equitably across the range of teaching assignments. There were concerns about inter-rater reliability as well as finding ways to differentiate professional development based on individual teacher needs as identified by the system. Teacher interviewees reported a concern among their peers about links to the teaching contract and potential ties between the results of teacher evaluation and personnel decision-making. One teacher stated, "It concerns me that it still seems like this is totally in the hands of the administration as to making sure that this is being equitably implemented."

Davies School District. Harmonious relations between committee members and a commitment to regular communication with educators across the district marked teacher

evaluation efforts in Davies School District. The committee was alert to accountability pressures as they designed a local plan and worked hard to allay teacher fears.

The community and its schools: Mixed school performance in a high-income community. Davies school district is located within 50 miles of one of Oregon's largest cities, and many residents commute to the city for work. The median household income in Davies is the highest of the three Oregon communities that are part of this study – significantly higher than the state figure of $50,229. The population has grown dramatically since 2000, by more than 25%. The crime rate in Davies is very low, lowest among the Oregon districts in the study by a significant amount. The percentage of Davies students who qualify for free or reduced-price lunch is more than 15 points below the state average, while per pupil spending in Davies is slightly below the state average and is the lowest among the three Oregon districts in this study.

Student outcomes in Davies are mixed. The district met 2013-2014 AMO targets in elementary and middle level reading, but failed to meet them in the remaining areas of elementary, middle, and high school mathematics and high school reading. The four-year graduation rate in this district is the lowest of the three Oregon districts in the study, slightly above the state figure of 69%. In 2013-14, more than 60% of Davies graduates went on to post-secondary education; this is above the state average and is the highest of the three Oregon districts in the study.

The average teacher salary in Davies is low, more than $5,000 below the state average, yet in 2013-14, the number of classes taught by “highly qualified” teachers was

above the state figure of 98.3%. There has been quite a bit of turnover in administration in Davies, with three new administrators coming to the district in the past few years.

Teacher evaluation decision-making: Collaboration, communication, trust. Davies assembled a teacher evaluation committee in the fall of 2012 in response to the passage of SB 290. District administration had begun learning about the policy and attending training sessions the previous spring. The committee included two representatives from the district office, two building principals, and two teachers from each building in the district. The teachers who served were volunteers, and there was an intentional effort to include active union members. The committee spent one year designing the local teacher evaluation system and one year piloting it, with the plan being fully implemented in the fall of 2014. During the first part of 2014-2015 when interviews were conducted, the committee was at work designing an approach to address the new student growth requirement and embedding the Oregon matrix into the local plan.

Davies developed its own teacher evaluation model, which they consider a hybrid. After a close review of teacher evaluation models from five Oregon school districts, the committee chose the plan being used in one of Oregon's largest districts as the basis for their local plan. To complete their design, Davies borrowed pieces from other state-approved rubrics, making adaptations to the base model to fit their much smaller district. In addition, the committee sought to align their local approach very tightly to the INTASC standards, referencing these standards continually throughout the design process and pulling the language of the standards directly into their evaluation protocols.

The density of policy collaboration on teacher evaluation was moderate to high in Davies. Multiple participants mentioned that they "learned together" and effectively made decisions as a team. Throughout the committee's work on teacher evaluation, their communication process involved the group's teacher representatives going out from the committee to each building to share information and receive feedback. It was important to the committee that district staff heard directly from teachers about teacher evaluation, as one of the district administrators explained:

> The teachers ran it, and I think that made a huge difference for us, was teachers hearing from other teachers that this is going to work, that this is going to be okay . . . Really letting our teachers be the trainers of other teachers.

It was clear that beyond teacher evaluation efforts, collaboration and trust are highly valued in this district, and multiple participants cited the good relationship that exists between administration and the teacher's association. The principal interviewee, who was relatively new to the district when the new teacher evaluation mandate came forward, expressed gratitude for the strong working relationship between the teachers union and district administration, having experienced contentious district-union relationships prior to coming to Davies:

> Our strength in this district is our people who prove that they care about kids. I also love the collaboration that I see with our teachers association here in this district that makes things easy to do because we all have the same goal, which is to serve our kids. And we don't get caught up in the typical district-union bickering that stalls education in my opinion.

The committee placed a strong value on developing an approach to teacher evaluation that would not be perceived as a "gotcha" system but, instead, a system aimed at professional growth. According to the principal interviewee:

> To me the biggest thing with regard to teacher evaluation is we need to – and I'm going to kind of get on my soap box here – we as principals need to be able to build relationships with teachers [so] that evaluation isn't seen as a ‘gotcha’ or a disciplinary thing . . . I want the evaluation portion of this to also be considered an opportunity for real and constructive feedback, both positive and challenges.

It should be noted that Davies’ work on teacher evaluation coincided with the initiation of data teams in the school district. This also informed the district’s approach. Following the pilot of the local model, committee members worked diligently to evaluate the model and make changes. A key emphasis for committee members was ensuring clarity in the teacher evaluation process and cleaning up areas of confusion.

Participants reported that the committee relied most on training sessions and resources from ODE, OEA, and the regional ESD. COSA, the Oregon School Personnel Association (OSPA), and the Oregon Education Investment Board (OEIB) were also cited. Despite the high level of involvement of teachers and principals in the local policy process, in the main, it was district office staff that attended state-level training sessions. In addition, these were the committee members who contacted state-level stakeholders for guidance; thus, interaction with state-level stakeholders was limited to a handful of individuals, resulting in a lower density of policy collaboration in this district compared to the other Oregon districts in the study. Though Davies was not a CLASS district,

committee members expressed appreciation for the work of CLASS districts and found the resources from these districts to be invaluable. The efforts of state-level stakeholders to work together on the rollout of the teacher evaluation policy were appreciated, and one participant praised ODE for effectively facilitating this collaboration across groups that historically do not always agree.

Perspectives on the policy: Concerns about policy context. Educators in Davies viewed the goals of Oregon's teacher evaluation policy to be improved teacher quality aimed at improved student learning. Participants also mentioned the goal of increased professional development and a focus on the use of data to improve instruction and student achievement. Key principles outlined in the district's teacher evaluation handbook illustrate an emphasis on establishing high expectations for teacher performance. A district administrator indicated his belief that a central goal of the policy is to place the proper emphasis on statewide testing in order to improve student achievement, and he tied this objective to professional learning:

> Ultimately the goals are to improve the achievement of students. To do so they're trying to create a system where teachers are rated or ranked by the performance of their students on specific tasks for different goals. Part of that is linked considerably to the testing system, the objective testing system, and there's two things in that: One is to make sure that we put significant emphasis on the testing system if people weren't already, and the other thing is to improve P.D. [professional development] for teachers, once again to help students score well on the assessments.

District administrators cited federal policy as the primary reason for the adoption of the teacher evaluation policy in Oregon. Others on the committee focused on issues of accountability in explaining why the policy was adopted, noting that Oregon's test scores and graduation rates have been very low. One principal stated:

> I think our testing scores are lagging, our grad rate is the worst in the country, and in order to have students meet outcomes, they have to have trained teachers and administrators that know what the heck they're doing to get them there.

A principal expressed the idea that the adoption of a state teacher evaluation policy was a "reactive move," typical in Oregon:

> In this state we've had a lot of reforms that are never allowed to go to fruition. It's kind of what the hot topic of the day is, or what the sexy device that the day is, and then that plays out for a couple years, and then something new comes along . . . We shift with the wind . . . Things aren't allowed to bear fruit; we cut the tree down too quickly.

In addition, participants questioned the federal waiver requirements, particularly the requirement that part of a teacher's evaluation be based on student growth scores, asserting that these requirements are not based on best practice. As a district administrator stated,

> I think what they are trying to do is improve instruction through these test scores . . . There's a real assumption that somehow people aren't trying hard enough, or they're not focused hard enough, and putting the score in place will make a

> difference . . . I haven't yet seen a theoretical model or any evidence that says it works.

Participants expressed major concerns over Oregon's matrix model, which was new at the time of the interviews. There was frustration over the fact that the model was added after the local plan was fully designed and piloted. In addition, participants were concerned about how the ratings data would be used and whether the media might publish teacher performance ratings. There was considerable annoyance about ODE's efforts on teacher evaluation, particularly regarding lack of clarity around the student growth requirement and how the transition from one state-mandated system of standardized testing to another would effect the evaluation system. At the same time, however, there was appreciation for ODE's willingness to respond to questions when the team needed information. The activities of ODE were described as being "frantic" efforts to secure the federal waiver. As a district administrator stated:

> The State Department [was] trying to get everything through to get the waiver passed . . . I don't think the rank and file took it as seriously because they have never taken NCLB as seriously as the bureaucratic apparatus in Salem . . . It was sort of trying to solve a political problem while they were working through it, and so the activity was hurried.

Participants seemed to understand that ODE was experiencing pressure from USDOE, yet frustration persisted.

The committee reported feeling good about its work and the quality of its teacher evaluation plan. They saw the new plan as an improvement on past practice, with the new

system being more effective in providing meaningful feedback to teachers and holding principals accountable for conducting regular evaluations of teaching staff. However, participants were concerned about how shifting state requirements might affect their work in a negative way.

Nilsenville School District. In a highly cooperative culture, educators in Nilsenville School District have capitalized on their work with the Chalkboard's CLASS Project to make teacher evaluation an important component of improving teaching and learning in their district. In addition, a strong relationship between the district and the teachers union has brought success to the effort.

The community and its schools: Performance concerns in a low-income district. Nilsenville School District is located in a small town where agriculture and timber have contributed to the history and economy of the community. The area attracts visitors seeking to enjoy the natural beauty of the region. The median household income in Nilsenville is the lowest of the three Oregon communities that are part of this study, and the crime rate is similar to that of Oregon at large.

Student performance on state mandated tests is a concern in Nilsenville. The district failed to meet target in either reading or mathematics at any level – elementary, middle, or high school. In OAKS testing, the percentage of Nilsenville students meeting or exceeding state standards was above the state average on less than half of the 10 exams. However, the 2012-2013 four-year graduation rate was higher than the state figure of 69%, and about the same percentage of Nilsenville graduates go on to post-secondary education as do graduates in the state at large.

In 2013-2014, the average salary of teachers in Nilsenville was more than $5,000 below the state average, while administrator salaries were close to the state average. Most of the district and building administrators in Nilsenville have been in the district for many years, and some have spent their entire careers here.

Teacher evaluation decision-making: Early policy success with Chalkboard support. As part of its work with Chalkboard, Nilsenville developed a teacher evaluation system along with new approaches to professional development and a robust mentoring and induction program. According to a district administrator interviewed for this study, the district received a total of between $800,000 and $900,000 over the three years of the CLASS Project (personal correspondence, June 30, 2015). The work with CLASS began some years before Oregon passed a teacher evaluation policy and applied for a federal ESEA waiver. Thus, the district's teacher evaluation system had been in place before other Oregon districts began their work on teacher evaluation in response to Oregon state law and federal policy. It should be noted that because time had passed since Nilsenville had completed its work with CLASS on developing a teacher evaluation system, participants' recollections of details were somewhat fuzzy. However, their memories of the "big picture" aspects of the work seemed clear, as did their impressions of the spirit of the work. In addition, because some of the educators who were originally involved in teacher evaluation decision-making had moved on from the district or had taken on different roles, study participants represented a mix of individuals – some of whom were part of the original decision-making process, and some of whom were not.

The effort to design a new teacher evaluation system in Nilsenville took place over two full school years, and the density of policy collaboration was high, with broad participation and cooperation and direct engagement with state-level stakeholders. Since no new policy requirements initiated the work, educators in this district based their decisions about teacher evaluation on research and best practice, not on meeting external demands. Participants described their work with CLASS as collaborative, and both administrators and teachers mentioned the district's history of a strong relationship between administration and the teachers union. In fact, it was clear that the intensely collaborative work associated with Nilsenville's CLASS project could not have been accomplished without this strong relationship. The team responsible for teacher evaluation design brought ideas to the broader staff for feedback dozens of times, both informally and through formal surveys; participants reported that among educators across the district, there was a lack of conflict surrounding the work. Of approximately 300 staff members, about 30 participated directly in the effort.

In general, participants were very positive about the work with CLASS and the benefits it brought to the district. A district administrator explained the effort:

> Part of that process was really looking at teacher effectiveness, knowing that teacher effectiveness has the greatest impact on student learning. And to do that they gave us an implementation grant where we were able to sit down – about 30 of our staff members together – and really start dreaming about what we thought would change education in [our district]. And this was a unique grant because they really let you do what you felt was right.

A participant who had been a union leader at the time when the CLASS project was initiated said that the union was quite resistant about becoming involved with CLASS when district leaders first proposed the project, but the two parties ultimately came to an understanding and invited Chalkboard into the district.

It is notable that teacher fear about developing a new system of teacher evaluation existed at the beginning of the district's work with CLASS, and, although participants reported their experience with CLASS to have been positive, it wasn't without its challenges. Teachers were resistant to the notion that CLASS represented business leaders coming in and telling schools what to do. As one interviewee explained:

> You have these people coming in from business and telling you . . . "This is really how you should be evaluating your employees," which that makes sense to me, but all that the people who were not on the design team heard was, "This is how you should be teaching." They're thinking, "What do you mean? You make windows! Why are you telling me how to teach?" When in reality that's not what they were doing. They were talking about evaluation and . . . how to retain people and train people, but teachers heard it as, "We're gonna teach you how to teach and do your job better cuz you're – you suck." That's what they – that doesn't sound very nice – but that's what they heard.

Now that it has been in place for several years, the teacher evaluation system is widely embraced. However, the relatively new requirement that student growth be included as a measure in teacher evaluation has caused teacher fear to resurface. (The teacher

evaluation system developed under CLASS did not include a measure of student growth or achievement.)

Participants gave much of the credit for the collaborative spirit in their work to the superintendent, who is trusted and considered a strong leader. He is well connected and participants expressed an appreciation for his ability to get “inside information” easily. Interviewees reported that Nilsenville is often ahead of other districts because of the superintendent’s connections. This perception is supported by the fact that the district has been looked to as an example of successful teacher evaluation practice; when Oregon's new teacher evaluation law passed, calls began to come in from other districts seeking Nilsenville’s help and advice. In addition, it was noted that Nilsenville's efforts in teacher evaluation influenced the shape of Oregon's teacher evaluation policy.

Adding the student growth measure to the existing system proved to be a difficult "sell" in Nilsenville. Part of the reason for this was that initial fears about teacher evaluation were connected to the suspicion that business leaders had a long-range goal of instituting a merit pay policy in Oregon. Also, through their work with the CLASS project, the district had been introduced to the idea of including a value-added modeling (VAM) component in their teacher evaluation system, but this idea was rejected. According to one participant, Chalkboard brought in an assessment expert, and it was a case where Chalkboard’s resources were not viewed kindly by district staff:

> They brought a guy out of the somewhere in the Midwest that was all about value-added models, and that was all he could focus on and all he could push for. He was a train wreck, I gotta tell ya . . . And so it was fortunate because you know

> what that did was that it galvanized our group that was working on this – and went wow, we are all together in that we are *not* doing this.

Thus, adding the student growth requirement to an existing system – one that was considered to be a success – proved to be a challenge. Also, participants noted that the frustration over adding the student growth measure was partly due to the fact that it was a top-down directive from the state. In their highly collaborative district where recent school improvement decisions had been made based on best practice, the notion of responding to an already rejected teacher evaluation measure was very frustrating.

In describing the external sources they relied on in developing their local approach to teacher evaluation, interviewees in Nilsenville praised CLASS and indicated that most of the additional resources that were made available to them as part of the CLASS project were highly valuable. One participant commented that this wealth of resources would never have been available to the district without their participation in the CLASS project. Additionally, the work with Chalkboard brought teacher evaluation leaders and the rest of district staff in contact with state-level stakeholders, broadening their understanding and ensuring that information was not filtered through district and/or building administration. Other external resources cited by educators from Nilsenville as being valuable supports during the design and implementation of their teacher evaluation system were ODE and COSA. Along with the Chalkboard Project, these two organizations were considered "best" resources by multiple participants.

After the new policy came out, local actors in Nilsenville sought guidance by attending trainings offered collaboratively by state-level stakeholders including ODE,

OEA, and COSA. In addition, representatives of these groups were invited to speak to Nilsenville educators about the new policy, and local actors appreciated the efforts of these groups to collaborate. Within the district, there was an intentional effort to communicate regularly about teacher evaluation design via a cadre of teachers on special assignment who were charged with providing in-person updates at school buildings across the district. For most teachers in the district, this was the means by which they learned about teacher evaluation planning.

Perspectives on the policy: An emphasis on accountability and quality. In response to questions about the goals of Oregon's teacher evaluation policy and the reasons for its adoption, participants in Nilsenville noted several key themes. Most prominent among the responses about the policy's goals were comments about teacher quality aimed at improved student learning. Also mentioned was holding teachers, administrators, and schools accountable. Interviewees cited federal requirements, the need to improve on past policy and practice, and poor test scores in their state as the primary motivating factors behind Oregon's new policy. The district's teacher evaluation handbook expresses a philosophy emphasizing diagnosis of individual performance needs to improve teacher performance and meet district goals for improved student learning. A teacher placed her comments about the policy's goals in the context of improving the profession:

> [The goal is] to give them kind of guidelines about where we want the profession at, the level we want. It's no longer just giving kids information, testing that they

have the information, and shoving them out the door. We have to do so much with kids nowadays that we need to have teachers at the right level.

Likely because their relatively new teacher evaluation system was considered a success, participants in Nilsenville were generally quite positive about Oregon's policy and related federal initiatives. However, as mentioned above, there were some concerns about recent policy developments. Concerning to participants was Oregon's new matrix for summative evaluation. As one interviewee stated, “There is no value in boiling it down to: You are one number . . . I think it's, again, the idea of a business model trying to put, you know, quantify it into one single thing.” In general, there was a sense in Nilsenville, as in the other Oregon districts, that policy requirements and related messages shifted frequently, causing confusion and additional work in local school districts. The principal interviewee expressed this frustration: "My anxieties have nothing to do with our district. It’s all from the state and how often things change."

Two aspects of the work of designing and implementing a new teacher evaluation system were cited as most challenging by Nilsenville participants: getting buy-in from teachers and ensuring fairness through inter-reliability training and calibration for principals. Regarding teacher buy-in, as mentioned earlier, concerns about a potential move to merit pay in Oregon predominated early conversations as the district was beginning its work with Chalkboard. It seemed that the only remaining concern about Nilsenville's local approach to teacher evaluation was the heavy workload for teachers and principals. This was especially acute for principals, since the district had also initiated an evaluation system for classified employees. However, it was clear that

participants considered the extra work to be worth the effort. As a principal interviewee stated:

> [The work] feels never ending, when you team it with all of the classified! . . . There's pre- and post- and it's just paperwork heavy . . . Even though I loathe the paperwork, I understand why and I agree with the theory. So I think that we're on the right track.

In their work to embed the new student growth requirement, the teacher evaluation committee tried to emphasize the use of locally developed assessments as much as possible. However, recent changes to the requirement forced the district to find ways to use state-mandated assessments for this component. As a result, the district intentionally de-emphasized the student growth component and the rating scale, placing the focus on improvement of practice, not scoring.

Participants viewed the state policy as having brought a focus to professional growth and support for teachers who struggle. In addition, they saw the policy as an improvement on past teacher evaluation practice and policy in their state. Regarding their local approach, educators in Nilsenville expressed appreciation for the clarity and consistency of their system; it was noted that the process of personnel decision-making became cleaner for both the district and the union under this system. The new teacher evaluation system and the accompanying changes that were made through the CLASS project resulted in a positive change in the district's culture, with an emphasis on teacher effectiveness as central to student success. A district administrator summed up the district's experience with its new teacher evaluation system:

> Really what we're tying to do is use [the evaluation rubric] to help us educate, train our teachers better. Evaluation is a necessary evil of that I guess, motivator in some cases. And we, in the course of the last five years, we have probably released more teachers than we ever have before because they were not highly effective. But that's all about our expectations . . . We don't tolerate mediocrity anymore.

Summary of the Three School District Cases

Data from interviews with local policy actors were analyzed from both the local school district and state (summary) perspectives, as articulated in Chapter Three. This section provides results from the summary analysis of interviews with the 13 participants from the three Oregon school districts that were part of this study. This section highlights results as they relate to the research questions and conceptual framework.

Interpretations of the policy's meaning. This study explored how local actors interpret teacher evaluation policies, and the first research sub-question addressed how they interpret the policy's meaning. Local actors responded to multiple questions that reveal how they interpreted the meaning of the policy, including queries about the goals of the policy, the reasons for its adoption, the most important requirements of the policy, and policy conflicts.

Local actors' understanding of policy goals: Professional growth for improved teaching and learning. The top two policy goals mentioned by local actors in Oregon were encouraging professional growth and ensuring teacher quality, the second of which is a policy aim not mentioned in interviews with state-level stakeholders. Both goals were

cited by educators from each of the three Oregon districts in the study, and by district administrators, principals, and teachers alike, though teachers were less likely than other participants to mention the goal of ensuring teacher quality. Two other goals arose in the analysis of local responses: improving student learning and standardizing teacher evaluation practice.

The main emphasis of those who named professional growth as a key policy goal was the importance of providing support to underperforming teachers. These participants noted that the system is designed to provide struggling teachers with help through targeted professional development. Those participants who cited ensuring teacher quality as a key policy goal accentuated the summative purpose of teacher evaluation, remarking about the need to raise expectations of teachers and noting the importance of Oregon's students having access to the best teachers. A large majority of local actors who mentioned ensuring teacher quality as a key policy goal were from Nilsenville School District, the one district in the study that had participated in the CLASS Project. Since CLASS focuses on strategies to support and improve teacher quality, this finding is not surprising. According to the CLASS page at the Chalkboard Project's website, CLASS is "built around four components linked to effective teaching: expanded career paths, effective performance evaluations, relevant professional development, and new compensation models" (Chalkboard Project, 2015a). Also, as mentioned earlier, Chalkboard's emphasis is on education reform, and teacher quality is central to recent reform efforts in the U.S.

The comments of local actors who stressed improving student learning in their explanation of the policy's goals focused on raising achievement scores or serving students well by providing a sound educational system, focusing on the summative purpose. Several local actors indicated that a central goal of the policy is to standardize teacher evaluation practice, both within districts and across the state of Oregon. This perspective reinforces data from interviews with both state and local policy actors who described a lack of consistency in teacher evaluation practice in the years preceding the adoption of the new policy.

Local actors' understanding of the reasons for the policy's adoption: Clarity about federal influence. When asked to explain their understanding of the reasons for the adoption of a new teacher evaluation policy in Oregon, participants expressed no hesitation or uncertainty, and their responses suggested a focus on the summative purpose of evaluation, especially accountability. The influence of federal policy was by far the most frequent response of local actors in answer to this question, and most participants attributed the policy's development to the sway of USDOE efforts to improve educator evaluation. Local actors who cited the federal influence did so matter-of-factly, with little or no elaboration. When asked to explain the reasons for the policy's adoption, one teacher simply responded, "Race to the Top. Federal policies. No Child Left Behind Waiver." A second, but far less frequent explanation regarding the reason for the policy's adoption was improving student achievement. Overall, participants seemed acutely aware that accountability pressures influenced the policy's development.

Local actors' views on the most important aspects of the policy: Feedback and fairness. When asked to comment on what they believed to be the most important aspects of Oregon's teacher evaluation policy, local actors identified three policy requirements: support and coaching from the principal, the use of multiple measures to evaluate teacher performance, and a standards-based system of teacher evaluation. Among these, support and coaching from the principal was cited most frequently. Those who felt that this was among the most important aspects of the policy stressed the relationship between the principal and the teacher and the opportunity for quality feedback on teaching performance.

Those participants who stressed the importance of multiple measures in the evaluation of teaching performance embraced the notion of setting goals for student learning as a valuable aspect of the student growth requirement. Only a few participants stated that they viewed the requirement for a standards-based system of teacher evaluation as being the most important aspect of the policy, and all were district administrators. It appears that to these district leaders, there is great value in having a system that bases its definition of proficiency on the standards of the profession and that the system can be applied consistently across a school district.

Policy conflict: Core values in question. To understand more fully local policy actors' interpretations of the meaning of the teacher evaluation policy, participants were asked about policy conflict, defined in Chapter One as conflict between the goals of the policy, or between the goals of the policy and the reasons for its adoption. In Oregon, almost all local actors indicated that they did indeed perceive conflict, and their responses

illustrate tension between policy mandates and the state's values, particularly regarding measures of student growth.

Almost all of the local actors who noted policy conflict discussed the student growth requirement and how the front-line realities of this requirement challenge core values about teaching and learning. Some local actors perceived conflict in the change Oregon has made in the student growth requirement to a more stringent approach. To these participants, this change, stimulated by federal waiver negotiations, brought a business mindset to the teacher evaluation policy and complicated decision-making about local systems. Other participants expressed the view that the use of standardized test results is an inappropriate approach, one that emphasizes accountability for accountability's sake over truly improving teaching and learning. Still others expressed concerns about the validity of this method, voicing fears about how the media might misuse the data produced under this approach, and questioning the "theory" that investments in standardized testing and efforts to tie results to teacher evaluation can actually improve instruction.

Influence of state-level organizations on local policy interpretations. This study's second sub-question addressed the process of policy interpretation, and the third sub-question concerned the influence of state-level policy actors on local policy activity. In addition, the conceptual framework portrays local policy actors as being influenced by state-level actors in their interpretations of the teacher evaluation policy. Thus, in this section we explore local actors' responses about whether the information and resources they received from sources outside their school district affected their view of the policy.

In Oregon, local actors acknowledged that their views on the policy's goals were shaped in part by state-level groups. About two-thirds of local participants were asked the question about whether they believed that the information, resources, and support they received from sources outside their district influenced their perspectives on the policy.[28] Several answered affirmatively, indicating the resources from external sources helped local teacher evaluation teams understand the policy's goals. A few others indicated that state-level stakeholders helped shape the design of local evaluation systems. However, a few local actors denied that their view of the policy was influenced by external groups or expressed uncertainty in response to the question.

State of Oregon Case

During this study's second phase, interviews were held with representatives of agencies or organizations that had been identified by local actors as reliable resources for information and support regarding the state's teacher evaluation policy. (For information on the characteristics of Phase II participants, see Chapter Three, Table 3, p. 93.) This section presents the interpretive lens of state-level policy actors on the study's research questions and conceptual framework. In addition, results regarding themes that arose over the course of the study, which were not specifically addressed by the study design, are described. These additional results include the perspectives of both local and state-level policy actors.

Interpretations of the policy's meaning. As stated above, this study's first research sub-question addresses interpretations of the teacher evaluation policy's

[28] This question was dropped from some interviews because of the nature of the responses to the questions leading up to it.

meaning. The conceptual framework portrays state-level policy actors as influencers of local policy interpretation, recognizing as well that these actors are interpreters themselves, who must seek to understand the teacher evaluation policy and its requirements in the context of state political culture and the broad policy ecology. State-level actors' interpretations of the policy's meaning can be understood through their responses to questions about the goals of the policy, the most important requirements of the policy, and policy conflicts.

State-level actors' understanding of policy goals: Teacher evaluation for professional growth. State-level policy actors were united in the view that Oregon's teacher evaluation policy is aimed at professional growth, which aligns with the formative purpose of evaluation.[29] The policy goal of supporting professional growth was established as a core value in the Oregon Framework, and each of the organizations that participated in this study was involved in developing this policy document. The framework identifies three main purposes of teacher evaluation, the first two of which reflect a focus on formative evaluation: strengthening educators' knowledge, skills and dispositions; strengthening support and professional growth opportunities for educators; and assisting districts to determine educator effectiveness for human resource decision-making. This third goal aligns to the summative purpose of evaluation. The framework also states that the ultimate goal of improving educator evaluation systems is to improve student learning (Oregon Department of Education, 2014a).

---

[29] In Missouri, both state and local actors referred to the policy's emphasis on "continuous improvement" to reflect the formative purpose, while in Oregon the language used was "professional growth." My analysis shows that the meaning is the same.

State-level actors' views on the most important policy requirements: A stress on systems and processes. To state policy actors in Oregon, the most important teacher evaluation policy requirements are the following: the requirement that districts evaluate teachers based on multiple measures of their performance, the requirement that the teacher evaluation system be based on standards of professional practice, and the requirement that administrators and teachers collaborate in the development of the local system. The first two of these requirements are among the five core elements for local evaluation systems outlined in the Oregon Framework, and two of the four participants referenced all five of those components in their response, indicating that all are important and serve as a foundation for effective evaluation practice. The third requirement emphasized as important by state-level stakeholders, that local systems be developed collaboratively, is a mandate in SB 290 and was noted as the most important by two of the four participants. The representative of Chalkboard stated, "I think the absolute most important one is that it is created and implemented collaboratively with equity of voice, across the school district." In general, responses of state-level participants to this question reflected pride in the Organ Framework and emphasized the importance of fairness and the promotion of excellent teaching practice through teacher evaluation systems.

Policy conflicts pose challenges. Though nearly all local actors in Oregon confirmed that they noted policy conflict related to Oregon's teacher evaluation policy, state-level policy actors were more mixed in their responses. Yet it is clear that policy conflicts brought challenges to policy interpretation and implementation planning.

When asked whether policy conflict was present in Oregon, the representative of the teachers association responded, “Oh yeah, definitely.” For this participant, there is conflict between Oregon’s attempts to create a teacher evaluation system that “enhances and promotes professional growth” and simultaneously seeking a waiver to get out of NCLB sanctions, a reality that this participant believes sets up decisions that don’t support the core values the state has established. This participant articulated serious concerns about the state’s move to growth score calculations based on standardized test scores, indicating that this approach not only goes against the state’s values, but also renders data that is useless for instructional decision-making at the classroom level:

> There have been things that we have, at a state level, agreed to that are not helpful, and will not be helpful, and do not uphold the goals or the values that were initially stated around these types of systems . . . Using student growth percentiles to measure an educator's effectiveness is a mathematically inaccurate task . . . Even if we could do it accurately or validly, it's really hard for many people to understand what that means . . . There's no way to draw the correlation of effect to those individual decisions, those split second decisions, those lesson-based decisions, even those unit-based decisions that you are making as you’re teaching and moving throughout the information that you are trying to impart on your kiddos . . . It becomes a meaningless exercise. It's a number, and it’s a sign. It's not a valid number because it can't actually measure your effect; it measures all effects on students.

Among the remaining state-level actors, one interviewee did not perceive any policy conflict, and the other two noted potential for conflict between the formative and summative purposes of evaluation. The ODE representative indicated that, depending on how the policy is interpreted, it is conceivable that some would see conflict in the fact that the policy is "written to support educators" yet requires that data from the system be used for personnel decision-making. The COSA representative indicated that there could be conflict between the policy goals of improving feedback to teachers and serving educational accountability demands, stating that the political discussion about teacher evaluation as an accountability tool is ahead of policy implementation, and that, thus, there is a gap between theory and practice.

Messaging and support from state-level organizations. Two of this study's research sub-questions have a connection to the work of state-level policy actors. These include the second sub-question, which addresses the process of interpreting teacher evaluation policy at the local level, and the third sub-question, which concerns state-level actors' influence on local policy interpretations. Also, the conceptual framework focuses on the policy interpretation process and the potential influence of state-level policy actors. State-level participants' responses to questions about the key messages their organizations communicated about the policy and about their efforts to support local educators in the rollout of the teacher evaluation policy help inform our understanding of how the policy was framed and how the policy interpretation phase was supported. Also included here is a discussion about collaborative efforts among state-level organizations.

Key policy message: "It's about helping teachers improve." Three of the four state-level policy actors answered the interview question about the key message their organizations communicated about teacher evaluation, and each of the three indicated that their main message was that teacher evaluation is about professional growth. The Chalkboard representative emphasized local collaboration in her response: "[Our message is] that it can become a professional growth tool when teachers are at the table in the design and implementation of it." The representative of ODE indicated that in its messaging, the department attempted to address apprehensions about teacher evaluation:

> We do try to let them know that this is about professional growth, that that's the most important part of this evaluation system, and to try to allay fears that some people automatically think that evaluation is about getting rid of bad teachers . . . That's not the intent. It's about helping teachers improve. And so professional growth is the key message.

According to the representative of OEA, the teachers association has attempted to communicate three messages about teacher evaluation, with the most important message being "that this is not so much an evaluation system but a professional growth system, which of course should and does have an evaluation component." OEA's other key messages are that teachers must "own" the opportunity to examine their impact on student learning, and that teacher evaluation should be evidence-based.

The representative of COSA was not asked directly about teacher evaluation messaging, and there is no formal statement about teacher evaluation available at the COSA website or in any of its publications. However, an analysis of the most recent

programs for the COSA annual conference, a four-day gathering held each June, shows that there were between one and three conference sessions on issues relating to educator evaluation in each of the five years following the passage of SB 290. Most of these sessions were aimed at helping school administrators understand how to implement the requirements of the policy, and no particular message about the intent of the policy is evident in the session descriptions or the conference themes (Confederation of Oregon School Administrators, 2013).

Support and resources provided by state-level organizations. In Oregon, collaboration across stakeholder groups marked state-level support for teacher evaluation. Though not without its challenges, the efforts of state-level stakeholders to join forces in providing training and guidance to districts proved fruitful. Throughout interviews with both local and state policy actors, the four organizations that participated in the second phase of the study were referenced as having been at the center of activity concerning policy guidance and implementation support for teacher evaluation. Of the four groups – ODE, OEA, Chalkboard, and COSA – the fourth group, COSA, played a less prominent role in these collaborative efforts. This was explained by the representative of OEA who stated that COSA took the lead on the other major policy initiative occurring simultaneously with teacher evaluation efforts, the work on implementing the Common Core State Standards; thus COSA's resources were stretched and they had less time to devote to the teacher evaluation work.

The four organizations whose representatives were interviewed for this study were involved in ongoing discussions about the waiver application, and this was the

group that designed the Oregon Matrix. Evidence that these partners took the lead in the educator evaluation work came in a press release in April 2014 announcing that ODE would be submitting the educator evaluation model to USDOE on May 1 of that year. The press release thanked the many educators who participated in the waiver work and the development of the Oregon Framework and was signed, "Oregon Educator Evaluation Partners, ODE | COSA | OEA | Chalkboard" (Chalkboard Project, 2015c). State-level policy actors were queried about the support and resources they provided to districts, as well as their communication strategies. What follows is a description of the work they did together and descriptions of each organization's independent work. Finally, a discussion of the impact of state-level collaboration on local educators is presented, which includes the perspectives of both local and state-level policy actors.

Collaborative work: A team of (sometimes) rivals. Beginning with the passage of SB 290, ODE began convening collaborative groups to work on aspects of moving the policy forward, and OEA, COSA, and Chalkboard were at the table. It should be noted that these organizations are sometimes opposed, butting heads on certain policy issues, yet on teacher evaluation they worked as a team. The first task was applying for the ESEA Flexibility Waiver, as described earlier. The group that was convened to work on the waiver application continued meeting throughout 2011-2012 to develop the Oregon Framework, which was endorsed by the SBE in June 2012. This document represents both policy guidance and an expression of the state's values around educator evaluation, articulated in the purposes and goals of the policy, the definition of teacher effectiveness included in the document, and in the decisions regarding details of the framework's

guidance not specifically expressed in policy. For example, the framework suggests a teacher evaluation system that begins with a self-reflection process, something that is required in neither SB 290 nor the federal waiver.

During 2012-2013, ODE partnered with OEA, Chalkboard, and the ESDs to offer day-long workshops throughout the state that outlined Oregon's vision for educator evaluation, the requirements of the policy, the use of multiple measures in educator evaluation, strategies for observation and feedback, and considerations for the evaluation of specialists (non-classroom teachers) (Oregon Department of Education, 2015a). Oregon's ESDs are defined in statute, and their purpose is to support the state and area school districts in meeting Oregon's education goals. The state's 19 ESDs are funded through a variety of sources, including property taxes, the school finance system, and grants or contracts. The ESDs are required to report to ODE, but they are not directed by the department (Oregon State Legislature, 2006). In the case of the regional training sessions on educator effectiveness, the ESDs served as hosts for the workshops; ESD personnel did not collaborate in the design or delivery of the training sessions. These workshops went beyond the delivery of information about policy requirements; educators who attended were offered information, resources, and consultation on best practices for key components of the Oregon Framework, and the primary intent of these workshops was to support districts as they developed local evaluation systems in time for the July 2013 deadline for submitting plans to ODE (Oregon Department of Education, 2015a).

In 2012-2013, when ODE piloted the Oregon Framework in 14 school districts, OEA and Chalkboard provided technical assistance to the pilot districts. Pilot districts

were not left alone to wrestle with policy demands; personnel from OEA and Chalkboard worked closely with local educators to assist in the pilot and to learn about issues and challenges so that they could inform other state-level policy actors about how the pilot was playing out on the front lines. The main focus of the pilot was to study the student growth component, so the technical assistance addressed areas such as developing, monitoring, and assessing student growth goals and calculating teacher performance ratings. According to the OEA representative, this represented the first time the teachers association had provided technical assistance in this manner.

Oregon Department of Education: Rushed yet resourceful. According to the ODE representative, in addition to the collaborative activities outlined above, ODE staff presented at conferences and workshops across the state, taking advantage of every opportunity to inform educators about the policy and its requirements. The participant stated, "We go to just pretty much every venue that our partner organizations have for professional development." The department developed an extensive set of tools, all available on ODE's website. This "toolkit" includes policy documents, timelines, research publications, webinars, tools such as goal setting forms and evaluation rubrics, guidance on local system design, model evaluation systems, and sample materials from Oregon school districts (Oregon Department of Education, 2015b). The ODE representative explained that in addition to communicating via its website, the department used its monthly newsletter to share information about educator evaluation. This participant also explained that the department utilized its stakeholder partners "to get the message out to their constituents."

Oregon Education Association: Far beyond lobbying. One of OEA's early efforts after the passage of SB 290 was the publication of a guidebook on teacher evaluation. This 54-page document outlined Oregon's policy and provided guidance for collaboration on system design at the local level. A CD-ROM containing resources on related issues such as effective collaboration, alternative compensation, and bargaining, as well as research summaries and sample evaluation frameworks accompanied the guidebook (Oregon Education Association, 2012). According to the OEA representative, the guidebook was one of the first published resources on educator evaluation in the state of Oregon.

The teachers association did quite a bit of training on teacher evaluation in addition to the collaborative training sessions on which they collaborated with ODE and Chalkboard. This included workshops that were done at the behest of local teachers unions, others where they presented at the invitation of school district administrators, and direct support services to individual school districts whereby OEA staff interacted one-on-one with local educators on policy activities. Between their collaborative work with state-level stakeholder organizations and the training sessions they did independently, at the time of the interview for this study, OEA Center for Great Public Schools (GPS Center) staff had provided training or direct support in 139 of Oregon's 197 school districts. Another important component of OEA's work was internal. Staff from the GPS Center ensured that OEA field staff – those staff members located around the state, working directly with local association members – had the materials and information they needed to support local educators in their teacher evaluation efforts. Especially important,

according to the OEA representative, was preparing field staff to help enforce the policy's requirement that teacher evaluation systems be collaboratively designed, with both administrators and teachers involved in the work.

OEA's communication strategy included posting a wide variety of materials at its website, including videos of the regional training sessions on which the organization collaborated with ODE and Chalkboard. Other means of communicating included utilizing field staff to communicate directly with local educators, sharing information via the organization's regular magazine and newsletter publications, and communicating through the association's elected leadership directly to elected leaders at the local level.

Coalition of Oregon School Administrators: Answers for administrators. The COSA representative explained that in addition to participating in the collaborative work described above, COSA offered informational sessions and updates across the state of Oregon at the 30-plus school administrator meetings it holds each year. At the time of the interview, the participant from COSA explained that such updates were continuing as policy implementation revealed additional questions. Often these sessions involved the participation of other stakeholder organizations such as the Oregon School Boards Association and OEIB (Confederation of Oregon School Administrators, 2013). According to the COSA representative, the group also did collaborative training sessions on teacher evaluation with school districts and with the Teacher Standards and Practices Commission, the state educator licensing agency. COSA's primary means of communicating about the policy was through its meetings and professional learning offerings.

Chalkboard Project: A focus on partner districts. As articulated above, the Chalkboard Project has been involved in supporting the development and implementation of educator evaluation systems since the mid-2000s, and this work helped pave the way for Oregon's teacher evaluation policy. After the policy's passage, along with the collaborative efforts described above, Chalkboard did training at the district level as well as "one-on-one coaching" with the 40-plus CLASS districts (personal correspondence, June 22, 2015). Primarily this coaching was with district teams, but there were times when Chalkboard worked one-on-one with individual teachers as well. In general, Chalkboard's work with CLASS districts involved deep interaction at the local level on teacher evaluation system design, along with consistent coaching and support over time. In addition, in partnership with OEA, Chalkboard provided technical assistance for the pilot of the Oregon Framework, as described above. Also, like OEA, Chalkboard published a booklet in 2012 that aided districts in understanding the new policy. *Oregon's Senate Bill 290: A Primer* was authored by a Chalkboard policy advisor along with an Oregon attorney with expertise in school law, and it provided policy interpretation and guidance to school district leaders regarding not only SB 290, but also Oregon's ESEA waiver (Hungerford & Dickson, 2012).

When asked about a communication strategy, the Chalkboard representative stated, "We usually let our districts speak for the great work that they're doing. We focus on results, so we measure impact through teacher perception, and student achievement, and achievement gap data, which is all across the board really positive." A review of Chalkboard's website reveals that such data is indeed used to promote the organization's

work. Data regarding the CLASS project districts (and the related TIF and Collaboration Grant sites) is summarized in documents that boast improvements in academic achievement, achievement gap reduction, graduation rates, and teaching conditions (Chalkboard Project, 2015a).

Impact of state-level collaboration: Intensity of effort with positive results. Although cooperation among state-level organizations is nothing new in Oregon, the intensity and intention with which education stakeholders collaborated on teacher evaluation went beyond what is typical, and it served to benefit school districts. Many local actors discussed the collaboration of state-level education stakeholders, offering positive remarks about how the collaboration was perceived. Some participants described the collaboration as an effort on the part of stakeholder groups to cooperate in interpreting the policy, something that was appreciated at the local level. It was acknowledged by some local actors that the state-level groups did not always agree, but that they worked out their differences in order to benefit Oregon school districts. A principal described a meeting he attended:

> We had two people from Chalkboard who came and talked with us, and then we also had our teachers union person from outside of our union – a state-level person – and those people sat down and really hashed out a lot of stuff in front of the group. And it was never contentious; it was just like, what exactly does this mean?

A few local actors commented that the efforts of state-level stakeholders to collaborate ensured that the voices of the various groups – teachers and administrators – were being

heard in the process and that local actors were getting a "balanced perspective." One teacher appreciated the involvement of the union:

> I remember when I went to the first few meetings I recognized a lot of my union friends that work for the [state] union . . . They were there as collaborators; they gave part of the trainings, as a matter of fact. And they were there to make sure that everybody understood what was being said – that what was being said was what was agreed on. So I truly feel that our union had a say in this.

Three of the four state-level participants (the representatives of ODE, OEA, and Chalkboard) were asked about the success of their collaborative efforts and whether they felt these efforts were beneficial to Oregon school districts, and all three affirmed that the efforts were successful. The ODE representative placed her positive remarks in the context of serving the department's need to get information, resources, and support to educators. To this participant, such collaboration ensures that "the information goes both ways" between the policy conversation at the state level and the constituents of the various stakeholder groups and, in addition, that the department can utilize the stakeholder groups to understand the needs of educators on the front lines. This participant went on to explain that the benefit to local educators was having their perspectives and needs represented in policy development, interpretation, and implementation, concluding, "You just need more voices in the process to represent the field." This emphasis on two-way communication between local and state-level policy actors was evident in interviews with all of the state-level participants in Oregon; it was

clear that a high value was placed on understanding how the policy was playing out on the front lines.

Asked if the collaboration was beneficial to educators at the local level, the Chalkboard representative responded, “Absolutely, yes.” This participant went on to explain that the collaboration built trust and helped to promote the “potential and power of a new evaluation system.” According to this participant, local educators often commented about the collaboration on evaluation forms they submitted at the end of training sessions presented by the partner organizations, indicating that they appreciated the efforts of state-level organizations to work together:

> When we took the end of training feedback – when you're really kind of wanting feedback about content and where to go next – we oftentimes got things like – under the question, “What was most successful?” – “Seeing all of you guys doing this together.”

The OEA representative discussed multiple ways in which the collaboration was beneficial to local educators, echoing the thoughts of the other participants. This participant also offered that the state-level collaboration served as a model for collaboration at the local level:

> It's been a good model for having issue driven collaboration – like, we don't have to agree on everything. We're going to put aside other things right now and focus on this and come to the table as equals. That's been a good model, I think, for districts to look at.

This participant went on to say that the modeling was especially helpful in districts where there were "more tentative or tenuous relationships between labor and management." In addition, this participant saw a benefit in how the organizations worked one-on-one with certain districts, trusting one another to have honest conversations with their respective constituents and "having each other's backs."

As did the local actors, state-level stakeholders acknowledged that their organizations didn't always agree; however, they worked through differences of opinion for the sake of the policy and its successful implementation. The OEA representative laughed when explaining that the organizations did not "get along all the time. We have our own opinions, but we really have made [collaboration] a value of the work that we have carried forth." In explaining why the effort was successful, the representative of Chalkboard stated, "We shared certain values, and where we disagreed we always came to consensus. So it was a great working relationship."

It seems that the development, interpretation, and implementation of Oregon's teacher evaluation policy brought a deeper level of involvement from state-level education stakeholders than past policy efforts had, particularly from the OEA. The ODE representative confirmed that, although stakeholder collaboration around education policy had been a "hallmark" in the state, the intensity with which stakeholders worked together to design and promote the teacher evaluation policy was something new. For the ODE representative, the increased involvement of OEA was especially helpful: "The fact that we were bringing OEA into the conversation right away at the beginning of this, I think has helped a lot. I would think that educators are more inclined to listen when OEA

is part of it." In describing the value of OEA's contributions, the ODE representative stated:

> Definitely OEA has been a tremendous partner . . . They have an arm of that organization which I'm sure you know by now is the center for great teachers, or great public schools. That part of the organization has really been dedicated to help teachers understand and grow. And they are all about improving practice and not just lobbying and advocacy for the organization – it's really about the professional learning for educators.

Another participant also called out the GPS Center as separate from the advocacy side of the union. This participant called the Center "amazing," and describing a good working relationship with their staff.

The representative of OEA, who works in the GPS Center, also commented on the more prominent role played by the association in the teacher evaluation work, placing OEA's efforts in the context of a "shift in unionism" which has been emphasized both by OEA and by its national affiliate, NEA. The participant described the move from a traditional union emphasis on issues such as bargaining and member rights to leadership in professional practice, where the stress is on "coming to the table" with solutions that support high standards of professional practice. The participant explained OEA's approach to the teacher evaluation work:

> We really wanted to be perceived as a partner and a collaborator, and not perceived pulling out a big gun or a big hammer. And that was a shock for some

people and a shift for them to think of the union on many sides of the aisle, to think of the union in that light.

Additional Results

In addition to results related directly to the research questions and the conceptual framework, three themes emerged during data collection. First, shifts in policy requirements and guidance, termed policy shifting in this paper, complicated policy interpretation and implementation. Secondly, participants expressed the belief that it was more difficult to interpret the teacher evaluation policy and plan for its implementation in certain types of districts. Finally, most participants expressed considerable optimism about the teacher evaluation policy.

Policy shifting: Fallout from back-and-forth with USDOE. Local policy actors in Oregon expressed frustration over the fact that the requirements of the teacher evaluation policy shifted over time, causing districts to revise their local designs in response. About half of local actors in Oregon commented on this, with district administrators most likely to register this concern. Most challenging and frustrating to teacher evaluation leaders at the local level was the fact that these changes came midstream, after plans had been developed and after buy-in had been established. Among local actors, there was broad understanding that shifts in the policy requirements were due to pressure from USDOE, and most participants seemed sympathetic with ODE staff and appreciative of their efforts to communicate changes as quickly as they could. In general, the tone of the concerns about policy shifting by local actors in Oregon was less critical than that of local actors in Missouri.

Nonetheless, fallout from shifts in policy requirements left local actors scrambling to rewrite plans, often in ways that left them uncomfortable and concerned about issues of validity and reliability inherent in the newly mandated measures. Two district administrators spoke at some length about the challenge of maintaining trust among staff members. One called the policy a "moving target," describing his district's experience with state and federal demands:

> Adding in new timelines by changing the requirements, by suddenly adding this matrix model at the end that none of us had planned for, and now we've got to weld it on. That was the hardest thing. When you're trying to build faith in people and you're trying to work with folks, trying to build an evaluation that's going to work well for everybody. That is the piece that's hard because it makes it look like for the rank and file that somehow you're lying to them.

State-level participants were aware of local concerns about changes in policy guidance and seemed as frustrated as some local educators. The ODE representative remarked that changes to the student growth requirement caused difficulties at the local level, stating:

> That hasn't played out real well with our folks in the field because it's a distinct change over time and that's not good . . . When we would come back and say, this is our policy, we do have to do this, it was a hard motivator.

Asked whether Oregon was watching the ESEA reauthorization discussions closely and tracking the potential that the approach to the student growth requirement might loosen,

the ODE representative responded, "Yes, and I'm hopeful that that will happen. That would be welcome in Oregon."[30]

Teacher evaluation considered more challenging for certain districts. Because several local actors in Oregon volunteered comments indicating that they believed it to be more difficult to develop local responses to the teacher evaluation policy in small or rural school districts, a question about whether the policy was more difficult to implement in small, rural, or high-poverty districts was added to the Phase II interview protocol. Among the issues raised by local actors were concerns about establishing and maintaining inter-rater reliability among evaluators in a small district. A district administrator explained the concern this way:

> I would say probably – and this is again because we're a small district – just being able to maintain calibration between the administrative team on what the observations look like . . . Because we are so small and because, frankly, we do tend to have some turnover because we're a little district and people kind of use it as a stepping stone to a bigger district. I think that's going to be a challenge for us, is just keeping that calibration going.

Another district administrator discussed an issue that had raised by others: the belief that, had the state provided more direction in the form of fully fleshed out teacher evaluation systems from which districts could choose – instead of simply recommending four rubrics that addressed only the observational requirement – local educators would have saved invaluable resources that were spent on doing the work locally, making it

[30] At the time of this interview, federal debates suggested that a reauthorized ESEA might not include student growth among required teacher evaluation measures. Presently, it appears that teacher evaluation could be removed entirely (Camera, 2015).

especially challenging for small districts. The district administrator articulated the concern as follows:

> We would have really liked to have seen the state say, okay here's SB 290, you have to do all these things, but here's a model to do it – here's one model, or maybe two. But what they gave us instead were five rubrics[31] we could pick from, but none of them were really an evaluation system . . . School districts have burned tons and tons of work hours on this, all doing the same work. If the state had had time and had been able to come up with a couple of models that were a full model – you know, here's an evaluation form, here's an observation form . . . I think that would have been really helpful, especially to smaller districts because we just don't have the man power to put in the amount of time on initiatives that we had to put into this. We don't have a curriculum director, or a school improvement person; we don't have any of that stuff.

Finally, a principal expressed the concern that there is a lack of equity when all districts are asked to meet the same requirements, regardless of their capacity to do so. His comment addresses district wealth and location:

> What works in a more affluent school isn't going to work in a more rural school just because of resources. And I have a unique perspective, having worked in very affluent schools, working in very lower-SES [socioeconomic status] schools, and this being my first time in a very rural school. So I've kind of gone the whole gamut, and it's interesting to see that all schools are expected to meet the same guidelines, but the equity is not where it needs to be in order to achieve it . . . Our

[31] The actual number of state-approved rubrics was four.

> kids are already behind the eight ball because they don't have the resources the other schools have, so it's no wonder.

The responses of state-level actors confirmed the perspectives of local actors; all of them expressed the belief that developing a local teacher evaluation system in response to Oregon's policy is more difficult in certain types of districts, and all cited small districts as being particularly challenged. In their responses, three of the four cited data on the large number of small districts in the state. The Chalkboard representative said:

> People rattle off this statistic all the time – I feel like we have 197 districts in Oregon and like 150 of them have under 2,500 students[32] . . . And the sheer capacity of anyone to do this work has been a huge challenge. They don't even have an H.R. [human resources] department or a curriculum department to have this land in.

Echoing the comment of the principal quoted above, the COSA representative expressed concern over the fact that every district must meet the same policy expectations, regardless of its capacity to do so. He summarized his comments by stating, "The practical reality is that size does matter." And in commenting about the lack of capacity for teacher evaluation efforts in both small and rural districts, the representative of ODE remarked about the difficulty of "few staff wearing many hats" in small school districts.

In addition to indicating that small districts are more challenged by the task of developing local responses to the policy, the representative of OEA noted the difficulty of doing the work in rural and rural remote districts. According to this participant, rural remote districts found it particularly challenging to meet the requirement that

[32] The actual number of districts with enrollments under 2,500 is 140. See Table 1, Chapter 3, p. 88.

administrators and teachers design the local teacher evaluation system collaboratively. Many of these districts are in parts of Oregon where there "tends to be more distrust of institutions in general, and so associations and unions fall under that label." Thus, the participant explained, collaborative work may go against the district's cultural norms, and cooperative relationships do not exist. According to this participant, in these districts, "forced collaboration didn't necessarily lead to a better product or a better design." This participant also noted that some districts with large populations of English learners, students in poverty, and/or nonwhite students had difficulty finding classroom observation rubrics that bore "an equity lens." The participant described two districts, one urban and one rural, where the local team modified one of Oregon's recommended rubrics in order to adequately address issues of cultural competence: "Those things became very difficult because you are kind of having to build something, having to take the time to have those conversations, which are meaningful and wonderful and rich, but it also takes time and it takes capacity."

Policy optimism: Confidence and hope. As mentioned in Chapter Four, policy optimism arose as a theme in the study's first phase, although there was no question about optimism on the interview protocol. Participants volunteered comments that illustrated their attitudes about the policy's chances of success. In Oregon, most local actors registered policy optimism, while the state actors, who were asked directly about how optimistic they were about the policy, all conveyed some degree of optimism.

Local actors in Oregon expressed optimism in one of three ways: they communicated confidence in their local teacher evaluation systems partly due to

satisfaction about the local process of developing the plan, they expressed gratification that the local plan had been received well across the district, or they conveyed hopefulness about the policy's potential for success.

Those participants who expressed confidence in the local plan seemed proud of the work they and their colleagues had done to develop the local teacher evaluation approach, and they viewed the resulting system as good for their respective school districts. A principal noted that the new policy brings important improvements:

> I think we're on the right path, and I think that making administrators more accountable for making sure that their staff members are doing well is a really good step in the right direction, which has never been part of it before. And I think that being more present in classrooms is very important.

A district administrator commented on teachers' reactions to the new system:

> They feel like [principals] can see what they're actually dealing with, and in most cases feel like they can help them become better teachers, and that's a huge step. So it hasn't been an adversarial what-are-you-doing-in-my-classroom type of situation.

There was a sense of accomplishment in the comments of local actors who happily reported that the newly developed teacher evaluation system had been received well by their colleagues. One principal stated, "I think every single person in my building that I have heard from views it as a positive." Participants whose optimistic comments were future oriented expressed the opinion that the policy will improve teaching and learning, as illustrated in this comment by a district administrator:

> Because of this new process and because we've had to go through such an extensive retraining about how we think about evaluation, I think the ultimate end result is going to be better for our teachers, better for our administrators, and certainly better for students.

Local actors characterized the new policy as a "step in the right direction," voicing hope that the policy will lead to success. It was indicated by some participants that, under the new policy, principals are optimistic that they can help teachers improve their instruction and that this will contribute to school improvement more broadly. As one teacher stated, "The goal of improving education is truly at the heart of what we're doing."

For state-level actors in Oregon, the answer to the interview question regarding their optimism about the policy's success was essentially, yes, but it depends. All four of the state actors stated that the policy holds great promise, but each mentioned concerns about school districts having the requisite support as policy implementation continues. State-level policy actors expressed concerns about both financial resources and professional development supports. The representative of COSA expressed no hesitation in indicating optimism "about compliance." However, this participant expressed the feeling that the policy's effectiveness will depend on resources, especially related to professional development for teachers, saying, "This policy comes up short" in providing resources to assist teachers, particularly those the system identifies as underperforming.

Both the representatives of Chalkboard and OEA expressed the concern that the teacher evaluation policy could lose momentum or be discarded if focus shifts to new policy reforms. The Chalkboard representative stated, "I get nervous that we are going to

be onto the next thing, and we haven't finished this thing," and the OEA representative explained similar concerns, concluding, "My hope is that this continues to be a backbone effort that is supported."

Despite these concerns, state-level actors in Oregon are hopeful that the state teacher evaluation policy – a policy each of them was involved in promoting – will meet its goals. The OEA representative's feelings capture the tone of state-level actors' hope for the policy's prospects: "I would say yes, I'm optimistic. I see the potential, and I've seen so much great success in certain efforts, like in certain districts, et cetera, that I can't help but feel optimistic around it."

Chapter Six

Findings and Discussion

This exploratory case study utilized an embedded multi-case design to investigate how policy actors interpret teacher evaluation policies in two U.S. states, Missouri and Oregon. The study was conducted in two phases: In the first phase, local policy actors from three small school districts in each state were interviewed about the state teacher evaluation policy and local efforts to meet the policy's requirements. In the study's second phase, state-level policy actors identified by local actors as valuable resources on the teacher evaluation policy were interviewed to gain a state perspective and to pursue themes that arose during the first phase of the study.

In this chapter, the research questions and conceptual framework are summarized. Next, the findings are presented, with those related to the research questions and conceptual framework coming first, followed by additional findings. Then, a discussion is provided, including implications for policymakers and practitioners as well as recommendations for future research.

Research Questions and Conceptual Framework

This study investigating the interpretation of teacher evaluation policies addressed the following research questions:

Central question:

How do local policy actors interpret teacher evaluation policies?

Sub-questions:

1. How do local policy actors interpret the *meaning* of teacher evaluation policies?
2. How do local policy actors experience the *process* of interpreting teacher evaluation policies?
3. How do state-level policy actors influence local responses to teacher evaluation policies?

The central research question can be addressed from two perspectives: *content* and *activity*, as indicated by the first two sub-questions. The third sub-question concerns the influence of state-level policy actors on local policy activity.

The study's conceptual framework places local policy actors within a broad policy ecology, nested within the influence of state-level policy intermediaries and the political culture of the state. (See Figure 1, Chapter Two, p. 79). This framework features the policy ecology (Weaver-Hightower, 2008) as a complex system where multiple actors, histories, and conditions bear on those involved in the process of policy interpretation. Next, the conceptual framework incorporates political culture – the political attitudes and behaviors associated with a particular geographical context – as an influence within the policy ecology that affects how policy actors at both the state and local levels interpret teacher evaluation policies (Elazar, 1970). In addition, this conceptual framework frames policy intermediaries – state-level policy actors – as both interpreters of policy and as influencers of local policy interpretation (Honig, 2004; Hamann & Lane, 2004; Hill, 2003). Finally, the conceptual framework places local policy actors, those involved in decision-making about the district teacher evaluation

system, at the center, influenced by all the other elements in the framework as they engage in the process of policy interpretation, and having the potential to influence directly the policy intermediaries with whom they interact. In addition, the framework acknowledges that influence flows both vertically and horizontally between and within the various actors and environmental factors involved in policy interpretation.

Findings

Interpretations of the policy's meaning. The first sub-question addressed local policy actors' interpretations of the meaning of new teacher evaluation policies, while the third sub-question addressed the influence of state-level policy actors on local policy activity, including their interpretation of the content of the policy. Certain findings are related to this question, having to do with perceptions of policy goals and requirements, and policy conflicts.

The formative purpose of teacher evaluation dominated policy actors' understanding of the policy's goals and most important requirements. Despite an increasingly public debate about teacher effectiveness, despite concerns about student growth measures, and despite suspicions that teacher evaluation policies are meant to punish rather than support teachers, local and state-level policy actors in this study embraced teacher evaluation as a means of professional growth. The power of the formative purpose dominated policy actors' discussions of policy goals and requirements, and they expressed optimism about the potential for the policy to improve teaching and learning. This finding supports research indicating that teachers view teacher evaluation positively and appreciate quality feedback that can help them grow professionally

(Kleinhenz & Ingvarson, 2004; Jiang, et al., 2015). However, it appears that for a minority of teachers, fears linger regarding teacher evaluation methods and the use of data from teacher evaluation systems – this despite the determined efforts of policy actors to ameliorate fear and skepticism.

Policy actors perceived conflict related to the two main purposes of teacher evaluation. This study confirms the tension between the formative and summative purposes of evaluation that was identified in the review of the literature on teacher evaluation. Both local and state-level actors in Missouri and Oregon interpreted this tension as a policy conflict.[33] At the local level, the conflict was manifested in the contrast between local actors' perception of the policy's goals, which was associated with the formative purpose, and their understanding of the reasons for its adoption, which was related to the summative purpose. For some, the policy's adoption was perceived as part of an agenda aimed at "getting teachers," and local policy actors were challenged to alter this perception as the policy process moved forward. At the state level, too, policy actors noted this policy conflict, with some articulating the difficulty of emphasizing professional growth within the teacher evaluation system while utilizing the data for personnel decision-making and accountability efforts. This finding about tension between teacher evaluation's formative and summative goals aligns to the finding from the literature review that new state teacher evaluation policies and the official guidance documents surrounding them center on formative evaluation, while the rhetoric

[33] As defined in Chapter One, policy conflict is conflict between a policy's goals, or between its goals and the reasons for its adoption.

concerning these policies relates to the summative purpose of evaluation, mainly accountability (Minnici, 2014; Harrison & Cohen-Vogel, 2012).

Although participants in this study identified the tension between the formative and summative purposes of teacher evaluation as a policy conflict, it appears that this conflict was not necessarily seen as a problem of practice. The literature review suggests that, in practice, school principals may experience role conflict due to the conflicting expectations of teacher evaluation, where, on one hand, the evaluator is expected to support and guide professional learning, and on the other, she is expected to make high stakes employment decisions and meet accountability demands (Biddle, 1986; Campbell, D.J. & Lee, 1988; Stiggins & Bridgeford, 1985; Popham, 1988, 2013). It is possible that, since full implementation of the teacher evaluation policy was just beginning at the time of the interviews for this study, issues related to role conflict may not yet have arisen.

Policy actors perceived conflict related to the student growth requirement. This study finds that the policy requirement to use measures of student growth in teacher evaluation caused policy conflict. Participants from both Missouri and Oregon interpreted conflict in the requirement itself and in how it was manifested in state policy and guidance. Study participants viewed the requirement as inherently problematic, complicating the process of policy interpretation and implementation planning. They articulated serious concerns regarding issues of validity in tying student performance to teacher evaluation, and they expressed frustration that the data generated for teacher evaluation purposes is useless for instructional decision-making at the classroom level. In addition, participants indicated conflict between the requirement and closely held

values regarding the nature of teaching and learning. They charged that the requirement inappropriately brings market-based values into educational decision-making, and they linked the requirement to a perceived over-investment in standardized testing. The concerns of educators about student growth requirements in teacher evaluation policies may reflect the finding from the literature review that a popular method for measuring a teacher's impact on student learning, value-added modeling (VAM), has faced considerable criticism and public controversy (Kupermintz, 2003; Baker et al., 2010; Corcoran & Goldhaber, 2013; American Statistical Association, 2014; Jiang et al., 2015; Scherrer, 2012). In addition, participants perceived a conflict between the student growth requirement and districts' abilities to implement the requirement effectively. In both states, shifts in the requirement that occurred after the policy was adopted, along with a lack of clarity from the SEAs about how the requirement was to be met, challenged implementation planning.

The process of policy interpretation. The second sub-question was closely related to the conceptual framework for this study, which depicts the *process* of interpreting teacher evaluation policy, while the third sub-question explored the influence of state-level policy actors on local policy activity. The findings presented in this section begin with local policy activity and move outward, addressing the influence of state-level policy actors on local policy activity, the political culture, and the broader policy ecology. It is acknowledged that some overlap exists; an attempt has been made to avoid redundancy wherever possible.

Local flexibility in teacher evaluation policy design led to intense efforts on the part of local policy actors. This study finds that state teacher evaluation policies affording local control in system design require a significant commitment by local policy actors. The new teacher evaluation policy found local policy actors engaged in detailed, exhaustive work over multiple years, and though the work was considered rewarding and fruitful, it did require a considerable investment of time and resources. With both states historically exhibiting a preference for localism in policy matters, most study participants embraced the policy's flexibility. However, a few local policy actors were frustrated by the need to invest so much time in local system design, wishing for more "direction" from the state. The study also illustrates that local processes of policy interpretation vary widely, yet commonalities exist. A notable distinction between districts was the degree to which administration controlled key decisions. In all three Missouri school districts, district and building administration held onto certain key decisions, ranging from near total exclusion of teachers from participation in one district to administration taking on the task of narrowing to two the teacher evaluation models for committee consideration in another. The effect in each Missouri district was the limiting, at least to some degree, of the involvement of teachers in policy interpretation and implementation planning. On the other hand, in Oregon, collaboration among teachers and administrators was the norm throughout the entire policy process. As an explanation for Oregon's highly cooperative effort, one can point to the requirement in the state's teacher evaluation law that administrators, teachers, and the exclusive bargaining representative of the teachers must work collaboratively on the local system, yet it appears that the state's moralistic political

culture, with its preference for broad participation in policy matters, may explain both the collaboration itself and the policy language encouraging it. Finally, it is noteworthy that in multiple districts across states, the process of policy interpretation involved a challenge: local policy actors found themselves having to employ extra efforts to address fear and skepticism about the teacher evaluation policy, as mentioned above.

The density of policy collaboration influenced the scope of policy interpretation at the local level.[34] This study illustrates that when there is broad participation and cooperation in policy interpretation across employee groups, when local policy actors participate frequently in policy activities, and when these actors are directly engaged in interactions with policy intermediaries, the policy interpretation phase is enriched and intensified, benefitting the policy process. In the districts in this study where the density of policy collaboration was moderate or high (one Missouri district and all three Oregon districts), local policy actors expressed a clear understanding of the policy, its requirements, and its intent; reported that they actively articulated their understanding to colleagues; developed ownership of policy meaning; and were more likely to exhibit optimism about the policy. Typically, districts where the density of collaboration was moderate or high involved a broader range of local policy actors in teacher evaluation decision-making, intentionally including teachers and their union representatives; convened teacher evaluation committees or teams frequently and involved all members in studying the policy and developing local approaches; offered training opportunities to teachers and building principals as well as district administrators; brought policy

[34] As defined in Chapter One, density of policy collaboration is the degree and frequency of participation in policy activities, the intensity of cooperation across groups, and the level at which local policy actors interact with state-level policy actors in relation to the policy.

intermediaries into the district for consultation and/or encouraged all local policy actors to interact directly with intermediaries; and were intentional in communicating about the policy to the broader staff, regularly soliciting feedback on local system design. Furthermore, in these districts, the heightened policy collaboration left policy actors feeling good about their own participation in the policy process and about the resulting local teacher evaluation system. Unfortunately, since interviews were conducted just as full implementation was taking place, there is no way to know whether density of policy collaboration had a positive impact on implementation of the teacher evaluation policy.

State-level stakeholders served as policy intermediaries, and their work benefitted local policy efforts. This study concludes that state-level policy actors served as policy intermediaries, interpreting the policy themselves and aiding local policy actors in policy interpretation. In addition, the study confirms Hamann & Lane's (2004) finding that state education agencies (SEAs) function as policy intermediaries. Moreover, the study finds that school districts benefitted from the work of state-level policy intermediaries.

To interpret the policy, local actors relied on policy intermediaries. This study reveals that local policy actors in both Missouri and Oregon relied on state-level policy intermediaries to aid them in interpreting the teacher evaluation policy. Policy intermediaries on whom teacher evaluation teams most relied were the SEAs in each state, state-level professional organizations representing teachers and school administrators, and organizations that provided direct support in designing local teacher evaluation systems. Further, in both states, local actors indicated that policy

intermediaries affected their view of the policy. In Missouri, intermediary organizations affected local actors' attitudes about policy implementation, while in Oregon they affected local actors' perceptions of the policy's goals. After the passage of the teacher evaluation policy, local policy actors were hungry for resources from trusted state-level organizations, and they utilized information made available to them in publications, websites, and training sessions to aid in their understanding of the policy and plan for implementation. These resources served as trail guides for the local journey toward implementation; in them local actors found details about policy expectations, research on best practices to aid local decision-making, optional "routes" for successfully meeting policy demands, and checklists to help chart progress. This study's findings on the influence of policy intermediaries confirm research indicating that intermediaries aid local actors in interpreting policy (Honig, 2004; Hill, 2003; Hamann & Lane, 2004).

There was a sharp contrast between the two states in the approach of policy intermediaries to the policy interpretation phase, and this influenced the nature of local actors' reliance on intermediary organizations. In Missouri, communication about the policy was unidirectional, with the SEA mainly transmitting information *to* districts. The SEA placed an emphasis on "getting the word out" regarding the policy and its requirements but did not engage proactively with other policy intermediaries to develop shared guidance and to learn about how policy interpretation and implementation planning were going at the local level. The state teachers association in Missouri that participated in this study shaped its communication strategy, not proactively on purposeful interactions with local policy actors, but reactively on questions the

organization received about the policy. However, the University of Missouri, which provided a teacher evaluation system districts could purchase, did display a commitment to two-way communication with the districts that were part of its network. It seems that as a result of the mainly unidirectional nature of communications between the state and local levels, local policy actors in Missouri positioned themselves primarily as consumers of information about the policy, rather than as active participants in policy discussions; local actors did not seem to view themselves as having the potential to influence interpretations of the policy's meaning beyond the boundaries of their own school districts.

By contrast, in Oregon, school districts benefitted from a deeper level of support from policy intermediaries, one that recognized the value of policy intermediaries actively *engaging with* and *listening to* local policy actors during the policy interpretation phase. As a result, local policy actors in Oregon seemed to perceive policy intermediaries as partners in policy interpretation, and they were aware that their interactions with intermediary organizations influenced policy interpretation in the state at large, not simply within their own school districts.

Messaging by policy intermediaries influenced local actors' policy interpretations. This study reveals that the key message of state-level policy intermediaries regarding the goals of the teacher evaluation policy influenced local actors' policy interpretations. In both Missouri and Oregon, policy intermediaries' primary message was that the teacher evaluation policy was intended to address the formative purpose of evaluation. This message had its intended effect: local actors in both

states emphasized the formative purpose as they described the policy's goals and its most important requirements. This finding supports research indicating that policy intermediaries help implementers learn about policy and interpret policy meaning (Hill, 2003; Hamann & Lane, 2004). However, in low-performing districts, the messaging from state-level stakeholders had less of an effect. Instead, local policy actors in these districts interpreted the policy as primarily aimed at accountability, the summative purpose. It appears that in low-performing districts, the message that prevailed had its source in the public narrative surrounding teacher evaluation, which frames low student achievement as the fault of teachers who don't – or won't – perform up to expectations and favors market-oriented approaches to school reform (Stern, 2013; Harrison & Cohen-Vogel, 2012; Kumashiro, 2012; Ravitch, 2012; Murphy, Hallinger & Heck, 2013).

Collaboration among policy intermediaries positively impacted local policy activity. This study finds that the efforts of policy intermediaries to collaborate on policy interpretation and implementation planning can have a positive impact at the local level. In Oregon, collaboration among state-level policy intermediaries aided local actors in understanding the policy and its requirements, ensured consistent guidance on policy interpretation, increased trust in the process of policy development and interpretation, provided a model for policy activities at the local level, and ensured members of various groups that their perspectives were represented as the policy was being rolled out. In addition, collaborations across intermediary organizations in Oregon ensured that local activities were understood and considered in state-level policy efforts, and policy intermediaries worked across boundaries to establish shared values, agree on negotiations

of policy meaning, and champion a policy some perceived as a threat to established policy positions, especially in the case of the state teachers union. Even in Missouri, where little state-level collaboration occurred, the limited efforts that did take place were appreciated by local educators who viewed the involvement of their state association representatives as giving them voice in the policy process. These findings support Honig's (2004) research demonstrating that intermediary organizations add value and that they expand their effectiveness by working with other intermediary groups to provide coordinated supports to implementers.

State political culture influenced policy activity at both the state and local levels. The oft-repeated finding that political culture explains differences in education policymaking in U.S. states is supported by this study, which reveals that the contrasting political cultures of Missouri, an individualistic state, and Oregon, a moralistic state, resulted in contrasts in policy activity surrounding teacher evaluation. In addition, this study confirms the finding of Louis and colleagues (2010) that the response of small school districts to policy initiatives is determined at least in part by state political culture. The contrasts in policy activity were found across the policy process. Concerning policy development, in Missouri, involvement in the drafting of legislation was restricted to policy elites, typical in an individualistic state, while in Oregon, policy development reflected the preference in moralistic states for the involvement of a broad range of stakeholders. Specific aspects of policy language were consistent with the political culture of the two states as well; in Missouri, the strong preference for local control caused revisions to bill language to secure flexibility in local system design, while in

Oregon, statutory language requiring consultation between administrators, teachers, and union representatives in teacher evaluation system design reflected the value placed on policy participation in moralistic state cultures. Responses to federal pressures related to the ESEA Flexibility Waiver differed in the two states, with Missouri policy actors rushing to meet federal demands in a process that limited participation from stakeholder groups and educators, and Oregon policy actors so committed to deliberation and consultation across groups that the state's waiver was placed in jeopardy. At the local level, as mentioned above, administrators in Missouri controlled teacher evaluation decision-making, while in Oregon, local policy activity involved district administrators, principals, and teachers equally – again displaying contrasts in the individualistic and moralistic cultures in terms of the value placed on participation in the policy process. Finally, local understandings of the state and national teacher evaluation policy context were deeper among local policy actors in Oregon, reflecting the tendency of citizens in moralistic states to actively engage in policy issues.

Additional findings. In addition to the above findings directly related to the research questions and the conceptual framework, this study reveals findings concerning policy shifting, challenges to policy activity in certain types of districts, and the new policy's relationship to past policy and practice.

Policy shifting related to federal demands was problematic. This study reveals that shifting teacher evaluation policy requirements challenged state and local policy actors in both Missouri and Oregon. Back-and-forth interactions between SEAs and USDOE as the states attempted to meet federal waiver demands caused states to alter

policy requirements and related guidance. This reality challenged state-level policy actors with additional work and complicated communications with school districts, and it caused particular consternation at the local level. Indeed, in both states local policy actors exhibited their frustration by describing the policy as a "moving target." In addition to causing difficulty for educators involved in designing local evaluation systems, this study finds that in both states shifting requirements affected trust – both between the SEA and school districts and between local teacher evaluation leaders and district staff.

Interpreting teacher evaluation policy and planning for implementation can be especially difficult in certain types of districts. This study finds that the task of interpreting teacher evaluation policy and planning for implementation can be especially challenging in small school districts where internal resources and external supports are lacking. Issues of limited capacity and lack of technical expertise challenge teacher evaluation policy enactment in small districts; both local and state policy actors noted that the complexity of the policy demanded skillsets that taxed local resources. In addition, it is perceived by some study participants that designing and implementing a local teacher evaluation system is particularly challenging in rural districts, remote districts, and districts with high concentrations of nonwhite students and English learners. This theme was more prevalent in Oregon than in Missouri.

The new policy was viewed as an improvement on past policy and practice. This study finds that in both Missouri and Oregon, state and local policy actors interpret the new teacher evaluation policy as an improvement on past policy and practice. Local actors viewed the old policies as outdated, characterized by formal observations that,

stereotypically, led some teachers to perform a “dog and pony show” when it came time for a visit from the principal. The new teacher evaluation policies were interpreted as being meaningful for teacher improvement, with their focus on dialogue and feedback about teaching and learning. In addition, the new policies are praised for bringing desired uniformity and consistency of compliance across districts. Furthermore, the study finds that optimism about the policy was high, with policy actors in both states expressing confidence about locally designed teacher evaluation systems and hope for the policy’s potential.

Discussion

This research distinguishes policy interpretation as a separate phase in the policymaking process, and it illustrates how policy actors at both state and local levels engage in concentrated efforts to interpret a policy and its requirements so that local implementation can lead to meaningful change. After a new policy is formally adopted and before full implementation takes effect, the policy activities in which state and local policy actors engage demand far more of them than simply following through with implementation. For policy actors in Missouri and Oregon, the process of interpreting the teacher evaluation policy involved learning about the policy, negotiating the meaning of policy requirements, filtering messages from the broad policy environment, collaborating across sometimes-opposed stakeholder groups, managing skepticism and fear, building trust, and tying teacher evaluation efforts to other recent educational policy changes. Under the conceptual framework for this study and in the findings presented here, policy actors are learners whose interpretations are influenced by a range of factors in the policy

ecology, including the history of teacher evaluation policy and practice, state political culture, the public narrative about teacher effectiveness, the teacher accountability movement, and media reports about teacher evaluation controversies. State-level stakeholder groups serve as policy intermediaries, aiding local actors in policy interpretation, and when intermediary organizations intentionally open channels of communication and engage with local policy actors during the interpretation phase, actors at the state and local levels influence one another.

Despite the political rhetoric invoking a "bad teacher" narrative that surrounds the current push for more and better teacher evaluation policies, policy actors embrace teacher evaluation for the promise of its formative purpose: professional growth. Even though policy actors are realistic in their understanding that the adoption of new teacher evaluation policies was brought on by a push for increased school and teacher accountability, policy actors – whether they be teachers, administrators, state-level representatives of these constituencies, SEA personnel, or representatives of support organizations – are optimistic about the potential of such policies to become meaningful systems for teacher growth. However, local educators remain wary about reform efforts that smack of a business mindset, and they fear the unfair use of test scores in teacher evaluation and the potential that unreliable methods may be utilized to remove teachers. It remains to be seen whether "the dysfunctional marriage of formative and summative teacher evaluation" (Popham, 1988) will provide a serious challenge to the success of new teacher evaluation policies that have been adopted across the U.S. in recent years.

Confirming research that describes the value of policy intermediaries, this study illustrates the importance of intermediaries as distinct actors in the policy process. Also, it appears that when policy intermediaries engage in substantive, long-term efforts to collaborate in the rollout of a new policy, school districts benefit. SEAs and other key stakeholder groups are portrayed in this study as promoters of the teacher evaluation policy, who framed key messages about the policy's intent and aided local actors in understanding the policy and its demands. Further, this study demonstrates that state-level intermediary organizations can be quite influential in shaping local policy interpretations, though in the case of teacher evaluation, the persistent public message emphasizing accountability for teachers and schools is difficult to overcome in districts challenged by low performance.

Implications for policymakers. For policymakers at both the state and federal levels, this study serves to illustrate the complexity faced by state and local policy actors in responding to external policy demands. In both Missouri and Oregon, efforts to make the teacher evaluation policy meaningful were fruitful, since optimism about the policy was generally high. However, frustrations with shifting requirements, tight timelines, and limited resources were evident. Back-and-forth negotiations between USDOE and SEAs placed SEA personnel and other state-level policy actors in a difficult position; in addition to working to meet federal demands under tight timelines, these policy actors bore the brunt of frustrations and anger from local educators who were forced repeatedly to revise local plans in response to changing requirements. In turn, local teacher evaluation leaders faced resentment and annoyance from their colleagues as they

communicated changes and reconvened teacher evaluation committees to modify local system designs. It is hoped that this study can help make policymakers aware of how shifts in policy requirements challenge policy actors, jeopardizing trust at all levels of the education policy system. Federal policymakers must also be aware that the policies they enact will be implemented in states with contrasting political cultures and that sensitivity to the peculiarities of how policy gets done in diverse locales may improve the odds of successful implementation.

Policymakers would be well served to attend to this study's finding that the requirement of most state teacher evaluation policies to include student growth data in teacher evaluation systems causes policy conflict, in terms concerns about the validity of applying such measures to teacher evaluation systems, challenges to deeply held values, and technical issues that complicate implementation planning. It is notable that frustrations about the use of student achievement data in teacher evaluation were shared across the various groups that participated in this study; teachers, district and building administrators, and state-level stakeholders alike found the requirement problematic. This policy conflict is complicated by the public controversy surrounding the inclusion of standardized testing data in evaluation systems, with multiple lawsuits pending and headlines reflecting the controversy having become commonplace. As Congress and the President prepare to reauthorize ESEA, facing proposals that eliminate the teacher evaluation requirement, federal policymakers will be forced to reconsider existing teacher evaluation mandates. In addition, state policymakers must reassess recently adopted state teacher evaluation laws, particularly their inclusion of student growth measures that

came, at least in part, in response to federal pressures. Along with the conclusions of Jiang and colleagues (2015), this study's finding that educators view new teacher evaluation policies positively, yet hold serious concerns about student growth measures, must be taken into account as teacher evaluation reform moves forward.

Implications for practitioners. This study can bring awareness to state-level practitioners, particularly those that serve as policy intermediaries, of their potential influence on local policy actors – to aid them, not only in understanding policy mandates, but also in interpreting policy goals and requirements, negotiating meaning, adapting the policy to align with other policy efforts, and championing a policy's potential for meaningful change at the local level. Further, this study's illustration of the power of collaboration across state-level stakeholder groups to provide a "web of intermediary supports" (Honig, 2004, p. 83) that build on each groups' strengths can be a lesson to SEAs, professional organizations, and other groups interested in ensuring that new education policies bring lasting change. Too, policy intermediaries must be aware that such collaboration may be difficult and may challenge organizational cultures, but the results of this study clearly indicate that local actors benefit when stakeholders go the extra mile to team with other groups for the good of schools.

For institutions, organizations, and individuals that provide for the preparation and professional development of educators, this study illustrates the need for teachers and school administrators to understand the policymaking process and their own potential for contributing to policy activities, particularly to the policy interpretation phase. In addition to understanding how education policymaking works, educators need to develop skills

that can help them become actively engaged in local policy efforts. In addition, teachers in particular – and the associations that represent them – should champion the roles teachers can play that go "far beyond lobbying." It is recommended that teachers associations provide professional development for their members on their potential to influence, not only policies that are related to traditional union concerns such as bargaining and benefits, but also policies that touch on the core professional issues of teaching and learning.

This research has implications for school districts as well. Those responsible for executing state and federal policies – particularly complex policies like teacher evaluation that relate to the core work of teaching and learning and require educators to meet new or significantly altered responsibilities – should understand that policy interpretation takes time and resources. Further, this study helps to illustrate the value of involving teachers and principals in local policy activity. Regarding teacher evaluation specifically, school district personnel should be aware of the tension between the formative and summative purposes of teacher evaluation, particularly of the potential that this tension may cause role conflict for principals. Training for principals that acknowledges potential role conflict may help to avert possible negative consequences. Also, the study's finding that teacher evaluation is valued by teachers and administrators as a means of supporting professional growth places a responsibility on school districts to capitalize on the opportunity offered by teacher evaluation reform.

Recommendations for future research. Few studies have investigated whether teacher evaluation achieves its formative purpose: the improvement of teaching practice.

Scholars have noted the difficulty of linking teacher evaluation to improvements in teaching (Colby et al, 2002), and many have articulated reasons for this difficulty, including lack of trust on the part of teachers, legal issues, lack of focus on professional learning, lack of validity in evaluation measures, and tension between evaluation's formative and summative goals (Duke, 1995; Sinnema & Robinson, 2007; Conley & Glasman; 2008). However, this study illustrates the potential – or at least the desire on the part of educators – for teacher evaluation to support improvements in teaching and learning. Further research on the formative purpose of teacher evaluation is needed to help policymakers and educators understand the particular aspects of teacher evaluation policies and systems that hold the most promise to impact teacher professional growth. Moreover, it is important to understand whether teachers, the subjects of the policy, view teacher evaluation as a tool for professional growth, as did educators in this study who expressed optimism about the policy's promise. Traditional means of measuring educational policy interventions – results of standardized tests – won't tell us whether teachers find new evaluation systems to have improved dialogue about teaching and learning in the school setting, whether they believe that the system has helped to address their professional development needs, and whether they credit the policy with improving their own teaching practice. Indeed, the voice of teachers must be included in the research agenda exploring the formative purpose of teacher evaluation policy.

With the recent expansion in teacher evaluation policies and the increased tension between evaluation's formative and summative goals, the potential role conflict experienced by school principals is an area worthy of further research. A limited number

of studies that specifically address this issue were discovered (Gmelch & Torelli, 1993; Clift, et al., 1992), and none published more recently than 20 years ago were found. Although this study did not identify the policy conflict between formative and summative evaluation as a problem of practice, it was viewed by study participants as a policy conflict, and stakeholders would benefit from a more thorough understanding of how principals experience the tension between teacher evaluation's formative and summative purposes and how this may impact the success of teacher evaluation systems. It will be important to know more about how principals deal with the public pressures toward accountability alongside pressures from within the profession to guide and support teachers in their professional growth.

Another area for research raised by this study is the under researched topic of collaboration across intermediary organizations. Honig's (2004) research found that such collaborative efforts can being value to policy efforts, yet little is known about what makes these collaborations effective. This study supports Honig's finding and illustrates how local educators benefit. Further research in this area may uncover particular strategies and approaches that can help to guide the work of intermediary organizations as well as the policymakers and school districts that rely on them.

Finally, this study introduces an important question for future research: Does the density of policy collaboration during the policy interpretation phase improve implementation? That is, is the efficacy of implementation enhanced when policy interpretation is characterized by broad participation and includes cooperation across critical stakeholder groups, when policy actors engage frequently in policy discussions,

and when local actors are given opportunities to interact directly with state-level policy intermediaries? As noted earlier, this research did not attempt to study implementation; indeed, it was conducted just as full implementation of new teacher evaluation policies was taking place in both Missouri and Oregon. Thus, although this study finds that in districts where the density of policy collaboration was high the scope of policy interpretation was enriched and intensified, there is no way to know its effect on policy implementation. At a time when states and districts across the U.S. are experiencing the first years under new teacher evaluation policies, the time is ripe to explore this question.

The policy focus on teacher evaluation as a means of improving public education finds multiple actors at all levels of the education policy system wrestling with policy demands. The intent of this study was to generate a firmer theoretical foundation for examining how policy actors respond to teacher evaluation reform and how that response is affected by state and federal policy conversations and other factors in the policy ecology. This study informs our understanding, not only of teacher evaluation policy, but also of policy interpretation as a separate phase in the policy process, shedding light on the actions and experiences of local and state-level policy actors in their quest to capitalize on new education policies to make meaningful changes in schools and classrooms. It is hoped that this study has given voice to the educators and stakeholders whose daily realities are interwoven with policy decisions that affect school cultures, classroom practices, and the professional lives of educators on the front lines.

References

ACT, Inc. (2014). 2014 National and state scores. Retrieved from http://www.act.org/newsroom/data/2014/states.html.lfonso, R. J., & Firth, G. R. (1990). Supervision: Needed research. *Journal of Curriculum & Supervision, 5*(2), 181-188. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=9604166946&site=ehost-live

American Federation of Teachers, AFL-CIO (2015). AFT-Oregon. Retrieved from http://or.aft.org

American Institutes for Research. (2014). AIR index: Teachers and evaluation reform. Retrieved from http://www.air.org/resource/air-index-teachers-and-evaluation-reform

American Statistical Association (April 8, 2014). *ASA statement on using value-added models for educational assessment*. Retrieved from http://www.amstat.org/policy/pdfs/asa_vam_statement.pdf

Amrein-Beardsley, A., Collins, C., Polasky, S. A., & Sloat, E. F. (2013). Value-added model (VAM) research for educational policy: Framing the issue. *Education Policy Analysis Archives*, *21*(4), n4. Retrieved from http://epaa.asu.edu/ojs/article/view/1311

Annie E. Casey Foundation. (2015). *Kids count 2015 data book*. Baltimore, MD: Annie E. Casey Foundation. Retrieved from http://www.aecf.org/resources/the-2015-kids-count-data-book/

Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., Rothstein, R., Shavelson, R. J., Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers.* (No. 278). Washington, DC: Economic Policy Institute. Retrieved from http://www.epi.org/publication/bp278/

Ball, S. J., Maguire, M., Braun, A., & Hoskins, K. (2011a). Policy actors: Doing policy work in schools. *Discourse: Studies in the Cultural Politics of Education*, *32*(4), 625-639. doi: 10.1080/01596306.2011.601565

Ball, S. J., Maguire, M., Braun, A., & Hoskins, K. (2011b). Policy subjects and policy actors in schools: Some necessary but insufficient analyses. *Discourse: Studies in the Cultural Politics of Education*, *32*(4), 611-624. doi: 10.1080/01596306.2011.601564

Ballotpedia (2015a). Missouri Commissioner of Education. Retrieved from http://ballotpedia.org/Missouri_Commissioner_of_Education

Ballotpedia (2015b). Missouri Teacher Performance Evaluation, Amendment 3. Retrieved from http://ballotpedia.org/Missouri_Teacher_Performance_Evaluation,_Amendment_3_(2014)

Barron, D. J., & Rakoff, T. D. (2013). In defense of big waiver. *Columbia Law Review, 113*(2), 265-345. Retrieved from http://www.jstor.org/stable/23479733

Beckman, D. (2015). Meet new Commissioner of Education Margie Vandeven. *Something Better*, pp. 8-9, 21. Jefferson City, MO: Missouri National Education Association. Retrieved from http://mydigimag.rrd.com/publication/?i=250620

Bergquist, G. (2014, May 4). Two petitions meet signature deadline. *Connect MidMissouri.com*. Retrieved from http://www.connectmidmissouri.com/news/story.aspx?id=1039650#.VUY8koHF-9j

Berry, K. S., & Herrington, C. D. (2011). States and their struggles with NCLB: Does the Obama blueprint get it right? *Peabody Journal of Education, 86*(3), 272-290. doi:10.1080/0161956X.2011.578982

Biddle, B. J. (1986). Recent development in role theory. *Annual Review of Sociology, 12*, 67-92. doi:10.2307/2083195

Bock, J. (2011, August 19). Troubled Missouri school districts fail to make gains. *St. Louis Post-Dispatch*. Retrieved from http://www.stltoday.com/news/local/education/troubled-missouri-school-districts-fail-to-make-gains/article_c0fbb122-20d2-57f2-af96-695babf1cd21.html

Bolin, F. S., & Panaritis, P. (1992). Searching for a common purpose: A perspective on the history of supervision. In C. D. Glickman (Ed.), *Supervision in transition: 1992*

yearbook of the association for supervision and curriculum development (pp. 30-43). Alexandria, VA: Association for Supervision and Curriculum Development.

Bolin, F. S. (1987). Perspectives and imperatives: On defining supervision. *Journal of Curriculum & Supervision, 2*(4), 368-380. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=48151120&site=ehost-live

Boston, C., Kovacs, K., Alexander, M. & Sitter, P. (2014, November 5). Two ballot measures pass statewide, and two are defeated. *Missourian*. Retrieved from http://www.columbiamissourian.com/a/181176/two-ballot-measures-pass-statewide-and-two-are-defeated/

Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (Third ed.). Los Angeles: Sage.

Brown, E. (2015, March 1). Teachers using courts to fight test-based evaluations. *The Washington Post*, p. A06.

Camera, L. (2015, August 5). Collision alert: House, Senate ESEA rewrites. *Education Week*, pp. 1, 21. Bethesda, MD: Editorial Projects in Education.

Campbell, R., Kyriakides, L., Muijs, R., & Robinson, W. (2003). Differential teacher effectiveness: Towards a model for research and teacher appraisal. *Oxford Review of Education, 29*(3), 347-362. doi:10.1080/0305498032000120292

Campbell, R., Kyriakides, L., Muijs, R., & Robinson, W. (2004). Effective teaching and values: Some implications for research and teacher appraisal. *Oxford Review of Education, 30*(4), 451-465. doi:10.1080/0305498042000303955

Campbell, D. J., & Lee, C. (1988). Self-appraisal in performance evaluation: Development versus evaluation. *The Academy of Management Review, 13*(2), 302-314. doi:10.2307/258579

Castillo, S. (2012, January 27). Oregon submits ESEA Flexibility Waiver. *Superintendent's update #354*. Retrieved from ww.ode.state.or.us/news/announcements/announcement.aspx?ID=7968&TypeID=4

Center on Great Teachers & Leaders. (2013). Databases on state teacher and principal evaluation policies. Retrieved from http://resource.tqsource.org/stateevaldb/

Chalkboard Project. (2015a). CLASS Project. Retrieved from http://chalkboardproject.org/what-we-do/class-project/

Chalkboard Project. (2015b). CLASS Project: Map of CLASS districts. Retrieved from http://chalkboardproject.org/wp-content/uploads/2010/12/CLASS_TIF_SDCF-Districts-in-Oregon-14.04.01.pdf

Chalkboard Project. (2015c). Press Room: An important message regarding Oregon's educator evaluation model [Press release]. Retrieved from

http://chalkboardproject.org/press-room/an-important-message-regarding-oregons-educator-evaluation-model/

Chalkboard Project, (2015d). What we do. Retrieved from http://chalkboardproject.org/what-we-do/

Chalkboard Project, (2015e). Who we are. Retrieved from http://chalkboardproject.org/who-we-are/

Clift, R., Johnson, M., Holland, P., & Veal, M. L. (1992). Developing the potential for collaborative school leadership. *American Educational Research Journal, 29*(4), 877-908. doi:10.2307/1163411

Coburn, C. E. (2005). The role of nonsystem actors in the relationship between policy and practice: The case of reading instruction in California. *Educational Evaluation and Policy Analysis, 27*(1), 23-52. Retrieved from http://www.jstor.org/stable/3699535

Cohen, D. K., & Barnes, C. A. (1993). Pedagogy and policy. In D. K. Cohen, M. W. McLaughlin & J. E. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 207-239). San Francisco: Jossey-Bass.

Colby, S. A., Bradshaw, L. K., & Joyner, R. L. (2002). Teacher evaluation: A review of the literature. Paper presented at the *Annual Meeting of the American Educational*

Research Association (New Orleans, LA, April 1-5, 2002), Retrieved from http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED464915

Cole, M. (2012, June 4). Oregon's state schools superintendent resigns, leaving the governor in charge of education. *The Oregonian*. Retrieved from http://www.oregonlive.com/politics/index.ssf/2012/06/oregons_schools_superintendent.html

Confederation of Oregon School Administrators. (2013). Professional learning. Retrieved from http://www.cosa.k12.or.us/professional-learning/events

Conley, S., & Glasman, N. S. (2008). Fear, the school organization, and teacher evaluation. *Educational Policy, 22*(1), 63-85. doi:10.1177/0895904807311297

Connell, R. (2013). The neoliberal cascade and education: An essay on the market agenda and its consequences. *Critical Studies in Education, 54*(2), 99-112. doi: 10.1080/17508487.2013.776990

Corcoran, S., & Goldhaber, D. (2013). Value added and its uses: Where you stand depends on where you sit. *Education*, *8*(3), 418-434. doi: 10.11 62/EDFP_a_001 04

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (Fourth ed.). Los Angeles: Sage.

Cruickshank, D. R., & Haefele, D. L. (1990). Research-based indicators: Is the glass half-full or half-empty? *Journal of Personnel Evaluation in Education, 4*(1), pp. 33-39. doi: 10.1007/BF00177128

Curators of the University of Missouri (2015a). Network for Educator Effectiveness: About NEE. Retrieved from http://nee.missouri.edu/about.aspx

Curators of the University of Missouri (2015b). Network for Educator Effectiveness: NEE Member Districts. Retrieved from http://nee.missouri.edu/Districts.aspx

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: ASCD.

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York; Oxford, OH: Teachers College Press; Learning Forward.

Darling-Hammond, L. (2014). What can PISA tell us about U.S. education policy? *New England Journal of Public Policy, 26*(1), Article 4. Retrieved from http://scholarworks.umb.edu/nejpp/vol26/iss1/4

Darling-Hammond, L., & Berry, B. (1986). *The evolution of teacher policy.* (No. ED 298 599). Washington, DC: Center for Policy Research in Education; RAND Corporation. Retrieved from http://files.eric.ed.gov/fulltext/ED298599.pdf

Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, 53*(3), 285-328. doi:10.3102/00346543053003285

DeBray, E., Scott, J., Lubienski, C. & Jabbar, H. (2014). Intermediary organizations in charter school policy coalitions: Evidence from New Orleans. *Educational Policy, 28*(2), 175-206. doi: 10.1177/0895904813514132

Delisle, D.S. (2013, August 14). Extension letter [Letter to Nancy Golden, Oregon Interim Chief Education Officer]. Retrieved from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/map/or.html

Devos, G., Ekholm, M., Kofod, K., Louis, K. S., Moos, L., Schratz, M., & van Velzen, B. (2012). Political culture and educational reform. In K.S. Louis & B. van Velzen (Eds.), *Educational policy in an international context: Political culture and its effects* (pp. 5-27). New York: Palgrave Macmillan.

Ding, C., & Sherman, H. (2006). Teaching effectiveness and student achievement: Examining the relationship. *Educational Research Quarterly, 29*(4), 39-49. Retrieved from http://web.a.ebscohost.com.ezp2.lib.umn.edu/ehost/detail/detail?vid=3&sid=f8163843-be12-464e-b27e-3e4c04822549%40sessionmgr4003&hid=4109&bdata=JkF1dGhUeXBlPWlwLHVpZCZzaXRlPWVob3N0LWxpdmU%3d#db=keh&AN=21110924

Doherty, K. M., & Jacobs, S. (2013). *Connect the dots: Using evaluations of teacher effectiveness to inform policy and practice.* Washington, DC: National Council on Teacher Quality. Retrieved from http://www.nctq.org/dmsStage/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report

Doore, A., Jorgensen, B., Doore, S., Mason, C. A. (2013). Teacher evaluation and professional growth systems: A national review of models, approaches, and challenges. Orono, ME: University of Maine. Retrieved from https://usm.maine.edu/sites/default/files/cepare/Teacher%20Eval%20Report%202013_04_B2.pdf

Doyle, W. (1977). Paradigms for research on teacher effectiveness. *Review of Research in Education, 5*(1), 163-198. Retrieved from http://www.jstor.org/stable/1167174

Duke, D. L. (1995). The move to reform teacher evaluation. In D. L. Duke (Ed.), *Teacher evaluation policy: From accountability to professional development*. Albany, NY: State University of New York Press.

Duncan, A. (2012, July 19). *Secretary's approval letter*. [Letter to Rudolph Crew, Oregon Chief Education Officer]. Retrieved from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/map/or.html

Dungca, N. (2012, August 16). Chalkboard Project awards $180,000 in teacher preparation and recruitment grants to Oregon districts. *The Oregonian*. Retrieved

from http://www.oregonlive.com/education/index.ssf/2012/08/chalkboard_project_awards_1800.html

Editorial Projects in Education. (2015a). Diplomas Count. Retrieved from http://www.edweek.org/ew/dc/index.html?intc=main-topnav

Editorial Projects in Education. (2015b). Quality Counts. Retrieved from http://www.edweek.org/ew/qc/index.html?intc=main-topnav

Elazar, D. J. (1970). *Cities of the prairie*. New York: Basic Books.

Ellett, C. D., Capie, W., & Johnson, C. E. (1980). Assessing teacher performance: Georgia's performance based certification project combines assessing teachers' performance with assisting in teachers' growth. *Educational Leadership, 38*(3), 219. Retrieved from http://web.b.ebscohost.com.ezp2.lib.umn.edu/ehost/detail/detail?vid=2&sid=0b9a7e84-3e8a-47ed-8f4a-609157cffccf%40sessionmgr112&hid=124&bdata=JkF1dGhUeXBlPWlwLHVpZCZzaXRlPWVob3N0LWxpdmU%3d#AN=9394794&db=aph

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education, 17*(1), 101-128. doi:10.1023/A:1025083214622

Elliot, K. (2011). Charlotte Danielson's framework for teaching sees record growth as states and districts redefine teacher evaluation. Retrieved from http://www.teachscape.com/about/press-and-news/2011/charlotte-danielsons-framework-for-teaching-sees-record-growth-as-states-and-districts-redefine-teacher-evaluation.html

Estes, K. (n.d.). *Missouri NEA: General presentation for new teachers*. Windsor C-1 NEA: Windsor, Missouri. Retrieved from http://wnea.weebly.com/downloads.html

Eye, G. G. (1975). Supervisory skills: The evolution of the art. *Journal of Educational Research, 69*(1), 14-19. Retrieved from http://www.jstor.org/stable/27536808

Fang, L. (2014, July/August). King Rex: A stock baron goes all in on Missouri politics. *Politico*. Retrieved from http://www.politico.com/magazine/story/2014/06/king-rex-sinquefield-108015.html#.VUZC6oHF-9j

Fehr, S. J. (2001). *The role of the educational supervisor in United States public schools from 1970 to 2000 as reflected in the supervision literature.* (Unpublished PhD). The Pennsylvania State University, State College, PA.

Feldvebel, A. M. (1980). Teacher evaluation: Ingredients of a credible model. *The Clearing House, 53*(9), 415-420. Retrieved from http://www.jstor.org/stable/30185365

Fenstermacher, G., & Richardson, V. (2005). On making determinations of quality in teaching. *The Teachers College Record, 107*(1), 186-213. Retrieved from http://www.tcrecord.org.ezp2.lib.umn.edu/library

Fine, J. (1997). School supervision for effective teaching in America. 1910-1930. *High School Journal, 80*(4), 40-54. Retrieved from http://www.jstor.org/stable/40364461

Firestone, W. A. (1989). Educational policy as an ecology of games. *Educational Researcher, 18*(7), 18–24. doi: 10.3102/0013189X018007018

Fowler, F. C. (2009). *Policy studies for educational leaders: An introduction* (Third ed.). Boston: Pearson.

Freeman, R. B. & Han, E. S. (2013). Public Sector Unionism Without Collective Bargaining. Paper presented at the *American Economic Association Annual Meeting, January 6, 2013*, San Diego, CA. Retrieved from http://dash.harvard.edu/bitstream/handle/1/12553710/82975005.pdf?sequence=1

Fuhrman, S. H., & Elmore, R. F. (1990). Understanding local control in the wake of state education reform. *Educational Evaluation and Policy Analysis, 12*(1), 82-96. Retrieved from http://www.jstor.org/stable/1163588

Furtwengler, C. B. (1991). A response to Milner. *Phi Delta Kappan, 72*(10), 785-787. Retrieved from http://www.jstor.org/stable/20404536

Furtwengler, C. B. (1995). State actions for personnel evaluation. *Education Policy Analysis Archives, 3*(1), 1-27. doi: http://dx.doi.org/10.14507/epaa.v3n4.1995

Gage, N. L. (1989). The paradigm wars and their aftermath: A "historical" sketch of research on teaching since 1989. *Educational Researcher, 18*(7), 4-10. Retrieved from http://www.jstor.org/stable/1177163

Gage, N. L., & Needels, M. C. (1989). Process-product research on teaching: A review of criticisms. *The Elementary School Journal, 89*(3), 253-300. Retrieved from http://www.jstor.org/stable/1001805

Gardner, D. P., Larsen, Y. W., Baker, W. O., & Campbell, A. (1983). *A Nation at Risk: The Imperative for Educational Reform.* Washington, DC: National Commission on Excellence in Education. Retrieved from http://files.eric.ed.gov/fulltext/ED226006.pdf

Garman, N. B., & Hazi, H. M. (1988). Teachers ask: Is there life after Madeline Hunter? *The Phi Delta Kappan, 69*(9), 669-672. doi:10.2307/20403745

Gibboney, R. A. (1987). A critique of Madeline Hunter's teaching model from Dewey's perspective. *Educational Leadership, 44*(5), 46. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=8718013&site=ehost-live

Glanz, J. (1977). Ahistoricism and school supervision: Notes towards a history. *Educational Leadership, 35*(2), 148-154. Retrieved from http://ascd.com/ASCD/pdf/journals/ed_lead/el_197711_glanz.pdf

Glanz, J. (1995). Exploring supervision history: An invitation and agenda. *Journal of Curriculum and Supervision, 10*(2), 95-113. Retrieved from http://web.b.ebscohost.com.ezp2.lib.umn.edu/ehost/detail/detail?vid=3&sid=422e5c86-dfdb-4989-91a9-bdd95bf567bb%40sessionmgr112&hid=115&bdata=JkF1dGhUeXBlPWlwLHVpZCZzaXRlPWVob3N0LWxpdmU%3d#AN=9610307881&db=aph

Glogowska, M., Young, P., & Lockyer, L. (2011). Propriety, process and purpose: Considerations of the use of the telephone interview method in an educational research study. *Higher Education, 62*(1), 17-26. doi:10.1007/s10734-010-9362-2

Gmelch, W. H., & Torelli, J. A. (1993). The association of role conflict and ambiguity with administrator stress and burnout. *Annual Meeting of the American Educational Research Association,* Atlanta, GA. Retrieved from http://eric.ed.gov/?id=ED359631

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf

Goldrick, L. (2002). *Issue brief: Improving teacher evaluation to improve teaching quality.* Washington, DC: National Governors Association Center for Best Practices. Retrieved from http://www.nga.org/files/live/sites/NGA/files/pdf/1202IMPROVINGTEACHEVAL.pdf

Guthrie, H. D., & Willower, D. J. (1973). The ceremonial congratulation: An analysis of principals' observation reports of classroom teaching. *The High School Journal, 56*(6), 284-290. Retrieved from http://www.jstor.org/stable/40366741

Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability, 26*(1), January 20, 2014. doi:10.1007/s11092-013-9179-5

Hamann, E. T., & Lane, B. (2004). The roles of state departments of education as policy intermediaries: Two cases. *Educational Policy*, *18*(3), 426-455. doi: 10.1177/0895904804265021

Hammond, B. (2013, July 16). Ex-education czar Rudy Crew racked up tons of travel, expenses, vacation, records show. The Oregonian. Retrieved from http://www.oregonlive.com/education/index.ssf/2013/07/ex-education_czar_rudy_crew_ra.html

Hammond, B. (2015a, April 8). Oregon schools chief Rob Saxton resigns. *The Oregonian*. Retrieved from

http://www.oregonlive.com/education/index.ssf/2015/04/oregon_schools_chief_rob_saxto_3.html

Hammond, B. (2015b, May 18). Oregon's next state schools chief: Salam Noor, former assistant superintendent of Oregon and Salem-Keizer schools. *The Oregonian*. Retrieved from http://www.oregonlive.com/education/index.ssf/2015/05/oregons_next_state_schools_chi.html#incart_story_package

Hanushek, E. A., Rivkin, S. G., Rothstein, R., & Podgursky, M. (2004). How to improve the supply of high-quality teachers. *Brookings Papers on Education Policy,* (7), 7-44. Retrieved from http://www.jstor.org/stable/20067265

Harrison, C., & Cohen-Vogel, L. (2012). The politics of teacher reform in Florida: Analyzing causal narratives surrounding state adoption of performance-based evaluations, performance pay, and tenure elimination. *Peabody Journal of Education, 87*(5), 517-534. Retrieved from http://www.jstor.org/stable/41725456

Hazi, H. M., & Rucinski, D. A. (2005). Refocusing on the ritual of teacher evaluation: Implications for the future of supervision. Paper presented at the *Council of Professors of Instructional Supervision Annual Meeting: Reflections of the Past, Visions for the Future, October 28-30, 2005,* University of Georgia, Athens Georgia. 154-168.

Hazi, H. M., & Rucinski, D. A. (2009). Teacher evaluation as a policy target for improved student learning: A fifty-state review of statute and regulatory action since NCLB. Retrieved from http://epaa.asu.edu/epaa/v17n5/

Higgins, J. (2015, September 19). After teacher strikes, all eyes are on Olympia. *The Seattle Times*. Retrieved from http://www.seattletimes.com/seattle-news/education/after-teacher-strikes-all-eyes-are-on-olympia/

Hill, H. C. (2000). *Implementation networks: Non-state resources for getting policy done* (Order No. 9963803). Available from Dissertations & Theses @ CIC Institutions; ProQuest Dissertations & Theses A&I. (304607894).

Hill, H. C. (2003). Understanding implementation: Street-level bureaucrats' resources for reform. *Journal of Public Administration Research and Theory, 13*(3), 265-282. doi: 10.1093/jopart/mug024

Hinchey, P. H. (2010). *Getting teacher assessment right: What policymakers can learn from research.* Boulder, CO: National Education Policy Center, University of Colorado at Boulder. Retrieved from http://nepc.colorado.edu/publication/getting-teacher-assessment-right

Hoffman, H. (2014, May 19). Chief education officer is highest-paid agency director. *Statesman Journal*. Retrieved from http://www.statesmanjournal.com/story/news/politics/state-workers/2014/05/18/chief-education-officer-highest-paid-agency-director/9195993/

Honig, M. I. (2004). The new middle management: Intermediary organizations in education policy implementation. *Educational Evaluation and Policy Analysis, 26*(1), 65-87. Retrieved from http://www.jstor.org/stable/3699504

Honig, M. (2006). Complexity and policy implementation. *New directions in education policy implementation: Confronting complexity*, 1-25. Retrieved from https://www.sunypress.edu/pdf/61303.pdf

Hughes, V.M. (2006). *Teacher Evaluation Practices and Teacher Job Satisfaction.* (Doctoral dissertation). Retrieved from https://mospace.umsystem.edu/xmlui/bitstream/handle/10355/4468/research.pdf?sequence=3

Hull, J. (2013). *Trends in teacher evaluation: How states are measuring teacher performance.* Alexandria, VA: National School Boards Association Center for Public Education. Retrieved from http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/Trends-in-Teacher-Evaluation-At-A-Glance/Trends-in-Teacher-Evaluation-Full-Report-PDF.pdf

Hungerford, N., & Dickson, K. (2012). *Oregon's Senate Bill 290: A primer*. Portland, OR: Chalkboard Project. Retrieved from http://chalkboardproject.org/wp-content/uploads/2012/08/FinalSB290PrimerExport8.8.pdf

Ingils, C. R. (1970). Let's do away with teacher evaluation. *The Clearing House, 44*(8), 451-456. Retrieved from http://www.jstor.org/stable/30184119

Iwanicki, E. F., & Rindone, D. A. (1985). Integrating professional development, teacher evaluation, and student learning: The evolution of teacher evaluation policy in Connecticut. In D. L. Duke (Ed.), *Teacher evaluation policy: From accountability to professional development* (pp. 65-98). Albany, NY: State University of New York Press.

Jaquiss, N. (2015, August 24). Education czar Nancy Golden's retirement raises questions about direction of K-12 reforms. *Willamette Week*. Retrieved from http://www.wweek.com/portland/blog-33640-education_czar_nancy.html

Jiang, J. Y., Sporte, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH students. *Educational Researcher*, *44*(2), 105-116. doi: 10.3102/0013189X15575517

Jones, A. S. (1972). A realistic approach to teacher evaluation. *The Clearing House, 46*(8), 474-481. Retrieved from http://www.jstor.org/stable/30188125

Klein, A. (2015, September 2). How long do Waiver states have to get teacher-evaluation systems in place? *Education Week*. Bethesda, MD: Editorial Projects in Education. Retrieved from http://blogs.edweek.org/edweek/campaign-k-12/2015/09/how_long_do_waiver_states_have.html

Kleinhenz, E., & Ingvarson, L. (2006). Teacher accountability in Australia: Current policies and practices and their relation to the improvement of teaching and learning. *Research Papers in Education, 19*(1), 31-49. doi:10.1080/0267152032000176963

Knapp, M. S. (2003). Professional development as a policy pathway. *Review of Research in Education, 113*(2), 109-157. Retrieved from http://www.jstor.org/stable/3568129

Kult, L. E. (1978). Improving teacher evaluations by principals. *The Clearing House, 52*(1), 17-21. Retrieved from http://www.jstor.org/stable/30185077

Kumashiro, K. K. (2012). *Bad teacher! How blaming teachers distorts the bigger picture*. New York: Teachers College Press.

Kumashiro, K. K. & Meiners, E. R. (2012). Flip the script. In G. M. Boldt & W. Ayers (Eds.), *Challenging the politics of the teacher accountability movement: Toward a more hopeful educational future* (pp. 12- 14). Bankstreet Occasional Papers 27. Retrieved from https://www.bankstreet.edu/occasional-paper-series/27

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis, 25*(3), 287-298. Retrieved from http://www.jstor.org/stable/3699496

Ladd, H. F., & Sorensen, L. C. (2014). Returns to teacher experience: Student achievement and motivation in middle school. *Center for Analysis of Longitudinal Data in Education Research Working Paper*, *112*. Retrieved from http://www.caldercenter.org/sites/default/files/WP-112_final.pdf

Larsen, M. A. (2005). A critical analysis of teacher evaluation policy trends. *Australian Journal of Education, 49*(3), 292-305. doi:10.1177/000494410504900306

Levinson, B. A., Sutton, M., & Winstead, T. (2009). Education policy as a practice of power theoretical tools, ethnographic methods, democratic options. *Educational Policy*, *23*(6), 767-795. doi:10.1177/0895904808320676

LexisNexis (2015). LexisNexis® Academic. Retrieved from http://www.lexisnexis.com/en-us/products/lexisnexis-academic.page

Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://www.gtlcenter.org/sites/default/files/docs/practicalGuide.pdf

Liu, P. (2015, June 16). Regents OKs new teacher evaluation system. Times Herald-Record. Retrieved from http://www.recordonline.com/article/20150616/NEWS/150619438

Looney, J. (2011). Developing High-Quality teachers: Teacher evaluation for improvement. *European Journal of Education, 46*(4), 440-455. Retrieved from http://www.jstor.org/stable/41343393

Louis, K. S. (1998). " A light feeling of chaos": Educational reform and policy in the United States. *Daedalus, 127*(4), 13-40. Retrieved from http://www.jstor.org/stable/20027521

Louis, K. S. (2012). E pluribus unum? Dissonance in US educational political culture. In K. S. Louis & B. van Velzen (Eds.), *Educational policy in an international context: Political culture and its effects* (pp. 5-27). New York: Palgrave Macmillan.

Louis, K. S., Gordon, M., Meath, J., & Thomas, E. (2009). The roots of difference in state educational policy. In B. C. Fusarelli & B. Cooper (Eds.), *The rising state: How state power is transforming our nation's schools* (pp. 153-176). Albany, NY: State University of New York Press.

Louis, K. S., Leithwood, K., Wahlstrom, K. L. & Anderson, S. E. (2010). Learning from leadership: Investigating the links to improved student learning. Minneapolis, MN: University of Minnesota. Retrieved from http://conservancy.umn.edu/bitstream/handle/11299/140885/Learning-from-Leadership_Final-Research-Report_July-2010.pdf?sequence=1

Louis, K. S., Thomas, E., & Anderson, S. (2010). How do states influence leadership in small districts? *Leadership and Policy in Schools, 9*(3), 328-366. doi:10.1080/15700761003731518

Louis, K. S., Thomas, E., Gordon, M. F., & Febey, K. S. (2008). State leadership for school improvement: An analysis of three states. *Educational Administration Quarterly, 44*(4), 562-592. doi:10.1177/0013161X08323858

Marsh, J. A., & Wohlstetter, P. (2013). Recent trends in intergovernmental relations: The resurgence of local actors in education policy. *Educational Researcher, 42*(5), 276-283. Retrieved from http://www.jstor.org/stable/23462393

Marshall, C., Mitchell, D., & Wirt, F. (1989). *Culture and education policy in the American states*. Bristol, PA: The Falmer Press.

Marshall, K. (2005). It's time to rethink teacher supervision and evaluation. *Phi Delta Kappan, 86*(10), 727-735. Retrieved from http://www.jstor.org/stable/20441897

Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: ASCD.

Mazzoni, T. (1994). State policymaking and school reform: Influences and influentials. In J. D. Scribner, & D. H. Layton (Eds.), *The study of educational politics* (pp. 53-73). Bristol, PA: The Falmer Press.

McDaniel, T. R. (1981). South Carolina's educator improvement act: Portent of the super school board? *Phi Delta Kappan, 63*(2), 117-119. Retrieved from http://www.jstor.org/stable/i20386185

McDonnell, L. M., & Elmore, R. F. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis, 9*(2), 133-152. Retrieved from http://www.jstor.org/stable/1163726

McGlone, A. (2013). Long-neglected law on teacher evaluations rises to forefront. *Hechinger Report,* (June 3), February 28, 2014. Retrieved from http://hechingerreport.org/content/long-neglected-law-on-teacher-evaluations-rises-to-forefront_12236/

McGuinn, P. (2012a). Fight club: Are advocacy organizations changing the politics of education? *Education Next, 12*(3). Retrieved from http://educationnext.org/fight-club

McGuinn, P. (2012b). Stimulating reform: Race to the Top, competitive grants and the Obama education agenda. *Educational Policy, 26*(1), 136-159. doi:10.1177/0895904811425911

McKenzie, P., & Santiago, P. (2005). *Teachers matter: Attracting, developing and retaining effective teachers.* Paris, France: Organisation for Economic Co-operation and Development. Retrieved from http://www.oecd.org/edu/school/attractingdevelopingandretainingeffectiveteachers-homepage.htm

Melnick, S. A., Shibles, M. R. & Gable, R. K. (1987). School District Size, Student Achievement and High School Course Offerings in Connecticut. *Research in Rural Education*, *4*(3), 119-23. Retrieved from http://jrre.vmhost.psu.edu/wp-content/uploads/2014/02/4-3_3.pdf

McNell, M. (2013, August 15). NCLB Waivers in Kansas, Oregon, Washington at 'high risk'. *Education Week*. Bethesda, MD: Editorial Projects in Education. Retrieved

from http://blogs.edweek.org/edweek/campaign-k-12/2013/08/nclb_waivers_in_kansas_washing.html

Mehta, J. (2013). How paradigms create politics: The transformation of American educational policy, 1980–2001. *American Educational Research Journal, 50*(2), 285-324. Retrieved from http://www.jstor.org/stable/23526143

Melton, K. (2011a, January 26). Oregon schools Superintendent Susan Castillo begins her third term with a new approach and an uncertain future. *The Oregonian*. Retrieved from http://www.oregonlive.com/education/index.ssf/2011/01/state_superintendent_susan_cas.html

Melton, K. (2011b, June 14). Oregon Senate passes bill to create statewide teaching standards. *The Oregonian*. Retrieved from http://www.oregonlive.com/politics/index.ssf/2011/06/oregon_senate_passes_bill_to_c_1.html

Melton, K. (2011c, February 15). Support grows for bill to appoint Oregon's school superintendent. *The Oregonian*. Retrieved from http://www.oregonlive.com/politics/index.ssf/2011/02/oregon_teachers_district_super.html

Merriam, S. B. (2009). Qualitative research: *A guide to design and implementation*. San Francisco: Jossey-Bass.

Meyer, H. H., Kay, E., & French Jr., John R. P. (1965). Split roles in performance appraisal. *Harvard Business Review, 43*(1), 123-129. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=3866369&site=ehost-live

Milanowski, A. T. (2005). Split roles in performance Evaluation – A field study involving new teachers. *Journal of Personnel Evaluation in Education, 18*(3), 153-169. doi:10.1007/s11092-006-9017-0

Millman, J. (1981). Chapter 1: Introduction. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 12-13). Beverly Hills, CA: Sage Publications.

Minnici, A. (2014). The mind shift in teacher evaluation: Where we stand – and where we need to go. *American Educator, 38*(1), 22-26. Retrieved from http://www.aft.org/periodical/american-educator/spring-2014/mind-shift-teacher-evaluation

Missouri Advisory Council of Certification for Educators, Mo. Rev. Stat. § 168.015.1 (2014). Retrieved from http://www.moga.mo.gov/mostatutes/stathtml/16800000151.html

Missouri Code of State Regulations: Rules of Department of Elementary and Secondary Education, § 5 CSR 20-400.375 (2014). Retrieved from https://www.sos.mo.gov/adrules/csr/current/5csr/5c20-400.pdf

Missouri Department of Elementary and Secondary Education (2015a). About the Commissioner of Education. Retrieved from https://dese.mo.gov/state-board-education/commissioner/about-commissioner-education

Missouri Department of Elementary and Secondary Education (2015b). About the State Board. Retrieved from http://dese.mo.gov/state-board-education/about-state-board

Missouri Department of Elementary and Secondary Education (2013a). Educator Evaluation System: Executive summary. Retrieved from https://dese.mo.gov/sites/default/files/eq-ees-executive-summary.pdf

Missouri Department of Elementary and Secondary Education (2012). Educator Evaluation System: Statutory Authority. Retrieved from https://dese.mo.gov/sites/default/files/eq-ees-statutory-authority.pdf

Missouri Department of Elementary and Secondary Education (2015c). Educator Growth Toolbox. Retrieved from https://dese.mo.gov/educator-quality/educator-growth-toolbox

Missouri Department of Elementary and Secondary Education (2015d). Educator Quality News and Updates. Retrieved from https://dese.mo.gov/educator-quality/news-and-updates

Missouri Department of Elementary and Secondary Education (2014). Missouri comprehensive data system. Retrieved from

http://mcds.dese.mo.gov/guidedinquiry/School%20Report%20Card/School%20Report%20Card.aspx

Missouri Department of Elementary and Secondary Education (2013b). Missouri Theory of Action. Retrieved from https://dese.mo.gov/sites/default/files/eq-ees-theory-of-action.pdf

Missouri Department of Elementary and Secondary Education (2015e). Overview. Retrieved from http://dese.mo.gov/state-board-education/overview

Missouri Department of Education (2015f). Senate Bill 320 Outstanding Schools Act. Retrieved from http://dese.mo.gov/governmental-affairs/legislation/previous-legislation/SB380

Missouri Department of Education (2015g). Top 20 by 20. Retrieved from http://dese.mo.gov/top-10-by-20

Missouri School Boards' Association (2010). Spring Meetings. Retrieved from http://www.msbanet.org/files/agreements/b3a0a5ba.pdf

Missouri State Teachers Association (2013a). 2013 Adopted resolutions: As adopted by the Missouri State Teachers Association Assembly of Delegates. Retrieved from http://www.msta.org/wp-content/uploads/2013/12/Adoped_Resolutions2013.pdf

Missouri State Teachers Association (2013b). *Articles of incorporation and bylaws of the Missouri State Teachers Association.* Columbia, MO: Missouri State Teachers

Association. Retrieved from http://www.msta.org/wp-content/uploads/2013/05/AriticlesBylaws_11-14-13_web.pdf

Missouri State Teachers Association. (2015). Our history. Retrieved from http://www.msta.org/about/our-history/

Mitra, D. L., Sanders, F. C. & Perkins, D. F. (2010). Providing spark and stability: The role of intermediary organizations in establishing school-based youth-adult partnerships. *Applied Developmental Science, 14*(2), 106-123. doi: 10.1080/10888691003738285

Munson, B. R. (1998). Peers observing peers: The better way to observe teachers. *Contemporary Education, 69*(2), 108-110.

Murhpy, J., Hallinger, P. & Heck, R. H. (2013). Leading via teacher evaluation: The Case of the missing clothes? *Educational Researcher, 42*(6), 349-354. doi: 10.3102/0013189X13499625

National Center for Education Statistics. (2015). The nation's report card. Retrieved from http://www.nationsreportcard.gov/reading_math_2013/#/

National Council of State Legislatures. (2015). Right-to-work resources. Retrieved from http://www.ncsl.org/research/labor-and-employment/right-to-work-laws-and-bills.aspx

National Council on Teacher Quality. (2014a). *2013 State teacher policy yearbook.* Washington, DC: National Council on Teacher Quality. Retrieved from http://www.nctq.org/dmsStage/2013_State_Teacher_Policy_Yearbook_National_Summary_NCTQ_Report

National Council on Teacher Quality. (2014b). Teacher evaluation timeline. Retrieved from http://www.nctq.org/dmsView/Evaluation_Timeline_Brief_Overview

National Education Association. (2015a). *Rankings of the States 2014 and Estimates of School Statistics 2015*. Retrieved from http://www.nea.org/home/rankings-and-estimates-2014-2015.html

National Education Association. (2015b). Teacher Quality. Retrieved from http://www.nea.org/home/20511.htm

Newton, R. R. (1980). Teacher evaluation: Focus on outcomes. *Peabody Journal of Education, 58*(1), 45-54. Retrieved from http://www.jstor.org/stable/1491522

Office of the Press Secretary, The White House. (2011). Press briefing by Press Secretary Jay Carney, Domestic Policy Council Director Melody Barnes, and Secretary of Education Arne Duncan, 8/8/2011. Retrieved from http://www.whitehouse.gov/the-press-office/2011/08/08/press-briefing-press-secretary-jay-carney-domestic-policy-council-direct

Orcher, L. T. (2005). *Conducting research: Social and behavioral science methods.* Glendale, CA: Pyrczak Publishing.

Oregon.gov. (2015). Oregon Transparency: Education Service Districts Transparency. Retrieved from http://www.oregon.gov/transparency/pages/esdtransparency.aspx

Oregon Department of Education (2015a). Educator effectiveness regional workshops 2012-13. Retrieved from http://www.ode.state.or.us/search/page/?=3872

Oregon Department of Education (2015b). Educator effectiveness toolkit. Retrieved from http://www.ode.state.or.us/search/page/?id=3759

Oregon Department of Education (2015c). ESEA Flexibility request: Renewal application submitted March 30, 2015. Retrieved from http://www.ode.state.or.us/search/page/?id=3475

Oregon Department of Education (2015d). Mission, values, and strategic plan. Retrieved from http://www.ode.state.or.us/search/page/?id=3933

Oregon Department of Education. (2014a). Oregon framework for teacher and administrator evaluation and support systems: Revised for 2014-2015. Salem, OR: Oregon Department of Education. Retrieved from http://www.ode.state.or.us/search/page/?id=3637

Oregon Department of Education. (2014b). Report card. Retrieved from http://www.ode.state.or.us/search/results/?id=116

Oregon Department of Education. (2012). *Report on the teacher/principal evaluation system baseline data collection for the State Fiscal Stabilization Fund*. Retrieved from http://www.ode.state.or.us/wma/teachlearn/educatoreffectiveness/quantitativeresult_educatorevaluationsystemsurvey_0830.pdf

Oregon Department of Education. (2011). State releases 2011 Adequate Yearly Progress results [Press release]. Retrieved from http://www.ode.state.or.us/wma/teachlearn/educatoreffectiveness/quantitativeresult_educatorevaluationsystemsurvey_0830.pdf http://www.ode.state.or.us/news/announcements/announcement.aspx?ID=7611&TypeID=5

Oregon Department of Education (2015e). Title II-A – Mentoring. Retrieved from http://www.ode.state.or.us/search/page/?id=2218

Oregon Education Association (2015). About OEA. Retrieved from https://www.oregoned.org/our-association/about-oea

Oregon Education Association (2012). *Teacher evaluation & support system guidebook: Tools for building an effective, collaboratively-designed evaluation system*. Portland, OR. Retrieved from http://www.oregoned.org/images/pages/OEA_GPS_EvaluationGuidebook_2012.pdf

Oregon Education Investment Board (2015a). Frequently asked questions. Retrieved from http://education.oregon.gov/Pages/Frequently-Asked-Questions.aspx

Oregon Education Investment Board (2015b). Network for Quality Teaching and Learning. Retrieved from http://education.oregon.gov/Pages/Network-for-Quality-Teaching-and-Learning.aspx

Oregon School Boards Association & Confederation of Oregon School Administrators (2011). *2011 OSBA/COSA Legislative Report.* Retrieved from http://www.osba.org/~/media/Files/Resources/Legislative/2011_Legislative_Report.pdf

Oregon School Boards Association & Confederation of Oregon School Administrators (2013). *2013 OSBA/COSA Legislative Report.* Retrieved from http://www.cosa.k12.or.us/downloads/legislative/2013_Leg_Report.pdf

Oregon Secretary of State. (2015). Oregon blue book: Public education in Oregon. Retrieved from http://bluebook.state.or.us/education/educationintro.htm

Oregon State Legislature. (2006). Efficiency study of the Oregon system of Education Service Districts. Austin, TX: Gibson Consulting Group. Retrieved from ww.ode.state.or.us/data/reports/docs/efficiency-study-of-the-oregon-system-of-education-service-districts-.pdf

Ornstein, A. C. (1995). Beyond effective teaching. *Peabody Journal of Education, 70*(2), 2-23. Retrieved from http://www.jstor.org/stable/1492844

Ovando, M. N. (2001). Teachers' perceptions of a learner-centered teacher evaluation system. *Journal of Personnel Evaluation in Education, 15*(3), 213-231. doi:10.1023/A:1012704809835

Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30*(2), 111-140. Retrieved from http://www.jstor.org/stable/30128057

Papay, J. P., & Kraft, M. A. (2015). Productivity returns to experience in the teacher labor market: Methodological challenges and new evidence on long-term career improvement. *Journal of Public Economics*. Retrieved from http://dx.doi:10.1016/j.jpubeco.2015.02.008

Pathe, S. & Choe, J. (2013, February 4). A brief overview of teacher evaluation controversies. *PBS NewsHour: The Rundown*. Retrieved from: http://www.pbs.org/newshour/rundown/teacher-evaluation-controversies/

Pearson, M. (2012, September 19). Win, losses and draws in Chicago school strike. *CNN*. Retrieved from http://www.cnn.com/2012/09/19/us/illinois-chicago-teachers-strike/

Pennington, K. (2014). *ESEA waivers and teacher-evaluation plans: State oversight of district-designed teacher-evaluation systems.* Washington, DC: Center for American Progress. Retrieved from https://www.americanprogress.org/issues/education/report/2014/05/07/89121/esea-waivers-and-teacher-evaluation-plans/

Peterson, K., & Kauchak, D. (1982). *Teacher evaluation: Perspectives, practices, and promises.* (No. ED 2333 996). Salt Lake City: Center for Educational Practice, Utah University. Retrieved from http://files.eric.ed.gov/fulltext/ED233996.pdf

Piro, J. S., Mullen, L. (2013). Outputs as educator effectiveness in the United States: Shifting towards political accountability. *International Journal of Educational Leadership Preparation, 8*(2), 59-77. Retrieved from http://files.eric.ed.gov/fulltext/EJ1016271.pdf

Podgursky, M., & Springer, M. G. (2007). Credentials versus performance: Review of the teacher performance pay research. *Peabody Journal of Education, 82*(4), 551-573. Retrieved from http://www.jstor.org/stable/25594760

Policy Innovators in Education Network (2015a). Who. Retrieved from http://www.pie-network.org/who

Policy Innovators in Education Network (2015b). Why. Retrieved from http://www.pie-network.org/why

Popham, W. J. (1972). California's precedent-setting teacher evaluation law. *Educational Researcher, 1*(7), 13-15. Retrieved from http://www.jstor.org/stable/1174524

Popham, W. J. (1988). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education, 1*(3), 269-273. doi:10.1007/BF00123822

Popham, W. J. (2013). *Evaluating America's teachers: Mission possible?*. Thousand Oaks, CA: Corwin.

Protect Our Local Schools (2014). About Us. Retrieved from http://www.protectourlocalschools.com/about/

Ravitch, D. (2012). The teacher accountability debate. In G. M. Boldt & W. Ayers (Eds.), *Challenging the politics of the teacher accountability movement: Toward a more hopeful educational future* (pp. 12- 14). Bankstreet Occasional Papers 27. Retrieved from https://www.bankstreet.edu/occasional-paper-series/27

Reed, P., Hylton, J., Cicirello, N. & Hall, S. (1988). *A Model plan for the supervision and evaluation of therapy services in educational settings* (ERIC Report No. Ed 318 191). Portland, OR: Child Development and Rehabilitation Center, Oregon Health Sciences University. Retrieved from http://files.eric.ed.gov/fulltext/ED318191.pdf

Reischman, C. (2014, September 16). Nicastro resignation ends contentious tenure. *The Missouri Times*. Retrieved from http://themissouritimes.com/13259/nicastro-resignation-ends-contentious-tenure/

Riley, C. (2014, December 18). Missouri's No. 2 education official promoted to No. 1. Springfield News Leader. Retrieved from http://www.news-leader.com/story/news/education/2014/12/17/missouris-education-official-promoted/20536297/

Robertson, J. (2014, December 17). Margie Vandeven is chosen from within to be the next Missouri education commissioner. *The Kansas City Star.* Retrieved from http://www.kansascity.com/news/local/article4557376.html

Rosenthal, A., & Fuhrman, S. (1981). *Legislative education leadership in the states.* Washington, DC: Institute for Educational Leadership. Retrieved from http://files.eric.ed.gov/fulltext/ED208549.pdf

Sawchuk, S. (2013, June 11). Race is on to ready teacher evaluations in New York City. *Education Week*. Bethesda, MD: Editorial Projects in Education. Retrieved from http://www.edweek.org/ew/articles/2013/06/12/35newyork.h32.html?qs=%22new+york%22+state+%22teacher+evaluation%22+cuomo+nysut

Sawchuck, S. (2015, September 11). Seattle teachers' strike encompasses range of issues. *Education Week*. Bethesda, MD: Editorial Projects in Education. Retrieved from

http://www.edweek.org/ew/articles/2015/09/16/seattle-teachers-strike-encompasses-range-of-issues.html?qs=seattle+teacher+strike+range

Scherrer, J. (2012). What's the value of VAM (Value-Added Modeling)? *Phi Delta Kappan*, *93*(8), 58. Retrieved from http://www.jstor.org/stable/23210376

Schweitzer, S. M. (1990). *Performance-based teacher evaluation in the state of Missouri: Teachers' perspectives* (Order No. 9129875). Available from ProQuest Dissertations & Theses A&I. (303869553). Retrieved from http://login.ezproxy.lib.umn.edu/login?url=http://search.proquest.com.ezp2.lib.umn.edu/docview/303869553?accountid=14586

Scriven, M. (1981). Summative teacher evaluation. In J. Millman (Ed.), *Handbook of teacher evaluation* (pp. 244-271). Beverly Hills, CA: Sage Publications.

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Research Quarterly, 61*(2), 294-308. doi:10.1177/1065912907313077

Sergiovanni, T. J. (1982). Introduction: Supervision and the improvement of instruction. In T. J. Sergiovanni (Ed.), *Supervision of teaching* (pp. vi-viii). Alexandria, VA: Association for Supervision and Curriculum Development. Retrieved from http://files.eric.ed.gov/fulltext/ED213075.pdf

Shapiro, J. (2014, May 4). Signatures filed for teacher tenure and early voting amendments to Missouri constitution. *The Kansas City Star*. Retrieved from http://www.kansascity.com/news/government-politics/article349791/Signatures-filed-for-teacher-tenure-and-early-voting-amendments-to-Missouri-constitution.html

Shiller, J. (2015). Speaking back to the neoliberal discourse on teaching: How U.S. teachers use social media to redefine teaching. *Critical Education, 6*(9). Retrieved from http://ojs.library.ubc.ca/index.php/criticaled/article/view/184931

Shimura, T. (2015, July 24). Lawsuit targets Saddleback Valley Unified, 12 others in state. *Orange County Register*. Retrieved from http://www.ocregister.com/articles/teachers-673379-lawsuit-districts.html

Sinnema, C. E., & Robinson, V. M. (2007). The leadership of teaching and learning: Implications for teacher evaluation. *Leadership and Policy in Schools, 6*(4), 319-343. doi:10.1080/15700760701431603

Sojourner, A. J., Mykerezi, E., & West, K. L. (2014). Teacher Pay Reform and Productivity Panel Data Evidence from Adoptions of Q-Comp in Minnesota. *Journal of Human Resources*, *49*(4), 945-981. Retrieved from http://jhr.uwpress.org.ezp2.lib.umn.edu/content/49/4/945.full.pdf

Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, *31*(2), 143-175. doi: 10.1080/002202799183205

Spillane, J.P., Reiser, B. J. & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research, *Review of Educational Research, 72*(3), p. 387-431. Retrieved from http://www.jstor.org/stable/3515992

Staiger, D. O., & Rockoff, J. E. (2010). Searching for effective teachers with imperfect information. *The Journal of Economic Perspectives, 24*(3), 97-117. doi:10.1257/jep.24.3.97

Stark, S., & Torrance, H. (2005). Case study. In B. Somekh, & C. Lewin (Eds.), *Research methods in the social sciences* (pp. 33-40). Los Angeles: Sage.

Stern, M. (2013). Bad teacher: What Race to the Top learned from the "race to the bottom". *Journal for Critical Education Policy Studies, 11*(3), April 15. Retrieved from http://www.jceps.com/archives/443

Stiggins, R. J., & Bridgeford, N. J. (1985). Performance assessment for teacher development. *Educational Evaluation and Policy Analysis, 7*(1), 85-97. Retrieved from http://www.jstor.org/stable/1164005

Stronge, J. H. (2007). *Qualities of effective teachers* (2nd Ed.). Alexandria, VA: ASCD.

Superfine, B. M. (2005). The politics of accountability: The rise and fall of goals 2000. *American Journal of Education, 112*(1), 10-43. doi:10.1086/444513

Taubman, P. (2012). Educational ~~reform~~ revolution. In G. M. Boldt & W. Ayers (Eds.), *Challenging the politics of the teacher accountability movement: Toward a more*

hopeful educational future (pp. 12- 14). Bankstreet Occasional Papers 27. Retrieved from https://www.bankstreet.edu/occasional-paper-series/27

Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review, 102*(7), 3628-3651. doi:10.1257/aer.102.7.3628

The Oregonian. (2015). 2011 Session: Senate Bill 290. Retrieved from http://gov.oregonlive.com/bill/2011/SB290/

Theriault, D. C. (2015, July 29). Ethics reforms after Kitzhaber: Going slow means getting it right, Oregon lawmakers argue. *The Oregonian*. Retrieved from http://www.oregonlive.com/politics/index.ssf/2015/07/ethics_reform_going_slow_means.html#incart_story_package

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education.* Washington, DC: Education Sector. Retrieved from https://www.sde.idaho.gov/site/charter_schools/research/Administrator/Rush%20to%20Judgement%20Teacher%20Evaluation.pdf

Tracy, S. J. (1995). How historical concepts of supervision relate to supervisory practices today. *The Clearing House, 68*(5), 320-325. Retrieved from http://www.jstor.org/stable/30189094

Tracy, S. J., & Smeaton, P. (1993). State-mandated assisting and assessing teachers: Levels of state control. *Journal of Personnel Evaluation in Education, 6*(3), 219-234. doi:10.1007/BF00125149

U.S. Census Bureau. (2015a). *2009-2013 American Community Survey 5-Year Estimates*. Washington, D.C.: U.S. Census Bureau. Retrieved from http://www.census.gov/data/developers/data-sets/acs-survey-5-year-data.html

U.S. Census Bureau. (2015b). *Public education finances: 2013*. Washington, D.C.: U.S. Census Bureau. Retrieved from http://www2.census.gov/govs/school/13f33pub.pdf

U.S. Department of Education. (2015a). ED data express. Retrieved from http://eddataexpress.ed.gov

U.S. Department of Education. (2015b). ESEA flexibility. Retrieved from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html

U.S. Department of Education. (2015c). NCLB / Proven methods. Retrieved from http://www2.ed.gov/nclb/methods/teachers/hqtflexibility.html

U.S. Department of Education. (2012). Oregon ESEA Flexibility Request (Original Submission). Retrieved from http://www2.ed.gov/policy/eseaflex/or.pdf

U.S. Department of Education. (2015d). Race to the Top Fund. Retrieved from http://www2.ed.gov/programs/racetothetop/index.html

U.S. Department of Education (2012, June 27). State of Missouri ESEA Flexibility Request (OMB Number: 1810-0708). Retrieved from http://www2.ed.gov/policy/eseaflex/approved-requests/mo.pdf

Ujifusa, A. (2013, November 21). Group seeking end to Missouri teacher tenure got help from K-12 chief Nicastro. *Education Week*. Bethesda, MD: Editorial Projects in Education. Retrieved from http://blogs.edweek.org/edweek/state_edwatch/2013/11/group_seeking_end_to_missouri_teacher_tenure_got_help_from_k-12_chief_nicastro.html

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122. Retrieved from http://www.jstor.org/stable/3516044

Weaver-Hightower, M.B. (2008). An ecology metaphor for educational policy analysis: A call to complexity. *Educational Researcher, 37*(3), 153-167. doi: 10.3102/0013189X08318050

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect.* Brooklyn, NY: The New Teacher Project. Retrieved from http://tntp.org/publications/view/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness

Westbury, M. (1988). The science and the art of teacher effectiveness: An examination of two research traditions. *Canadian Journal of Education, 13*(1), 138-161. Retrieved from http://www.jstor.org/stable/1495173

Whalen, A. (2015, July 23). *Renewal approval letter* [Letter to Salem Noor, Oregon Deputy Superintendent of Public Instruction]. Retrieved from http://www2.ed.gov/policy/eseaflex/secretary-letters/orrenewalltr2015.pdf

Wilson, S. M., Darling-Hammond, L., & Berry, B. (2001). *A case of successful teaching policy: Connecticut's long-term efforts to improve teaching and learning.* (No. R-01-2). Seattle, WA: Center for the Study of Teaching and Policy, University of Washington. Retrieved from http://files.eric.ed.gov/fulltext/ED475448.pdf

Wilson, V., & Edwards, L. (2001). *Telephone interviewing in educational settings*. (No. 84). Edinburgh: Scottish Council for Research in Education. Retrieved from http://www.eep.ac.uk/dnn2/Organizations/SCRE/tabid/86/Default.aspx

Winkler, A. M., Scull, J. & Zeehandelaar, D. (2012). *How strong are U.S. teacher unions? A state-by-state comparison*. Washington, D. C.: Thomas B. Fordham Institute & Education Reform Now. Retrieved from http://edexcellence.net/publications/how-strong-are-us-teacher-unions.html

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1985). Teacher evaluation: A study of effective practices. *The Elementary School Journal,* 61-121. Retrieved from http://www.jstor.org/stable/1001217

Wiswall, M. (2013). The dynamics of teacher quality. *Journal of Public Economics*, *100*(2013), 61-78. Retrieved from http://dx.doi.org/10.1016/j.jpubeco.2013.01.006

World Media Group, LLC. (2015). USA.com U.S. crime index state rank. Retrieved from http://www.usa.com/rank/us--crime-index--state-rank.htm

Worner, W. M. (1981). Small School Districts and the High Cost of Energy. *Journal of Education Finance, 6*(3), 297-309. Retrieved from http://www.jstor.org/stable/40703283

Yeh, S.S. (2009). The cost-effectiveness of raising teacher quality. *Educational Research Review, 4*(3), 220-232. doi:10.1016/j.edurev.2008.06.002

Yin, R. K. (2014). *Case study research: Design and methods* (Fifth ed.). Los Angeles: Sage.

Yokley, E. (2015, February 11). Unions target of House bills. *The Joplin Globe*. Retrieved from http://www.joplinglobe.com/news/local_news/unions-target-of-house-bills/article_4af2890a-2962-514b-907b-c40a278c05f2.html

Yudin, M. (2012, April 17). *Letter regarding peer review feedback* [Letter to Chris L. Nicastro, Missouri Commissioner of Education]. Retrieved from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/map/mo.html

Yudin, M. (2012, April 17). *Letter regarding peer review feedback* [Letter to Susan Castillo, Oregon State Superintendent]. Retrieved from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/map/or.html

Appendix A

Status of U.S. States Regarding ESEA Flexibility, Race to the Top Awards and State Teacher Evaluation Policy Implementation

State	Approved for ESEA Flexibility[a]	Awarded Race to the Top Grant[b]	Full implementation of teacher evaluation policy[c, d]
Alabama	✓		2015-2016
Alaska	✓		2014-2015
Arizona	✓	✓	2012-2013
Arkansas	✓		2014-2015
California			1971 law still in place
Colorado	✓	✓	2013-2014
Connecticut	✓		2013-2014
Dist. of Columbia	✓	✓	2014-2015
Delaware	✓	✓	2012-2013
Florida	✓	✓	2011-2012
Georgia	✓	✓	2014-2015
Hawaii	✓	✓	2013-2014
Idaho	✓		2014-2015
Illinois	✓	✓	2016-2017
Indiana	✓		2012-2013
Iowa	Under review		No state authority to enforce policy
Kansas	✓		2014-2015
Kentucky	✓	✓	2014-2015
Louisiana	✓	✓	2012-2013
Maine	✓		2015-2016
Maryland	✓	✓	2013-2014
Massachusetts	✓	✓	2014-2015
Michigan	✓		2015-2016
Minnesota	✓		2014-2015
Mississippi	✓		2014-2015
Missouri	✓		2014-2015
Montana			2013-2014
Nebraska			Voluntary compliance

Continued on next page

State	Approved for ESEA Flexibility[a]	Awarded Race to the Top Grant[b]	Full implementation of teacher evaluation policy[c, d]
Nevada	✓		2014-2015
New Hampshire	✓		2014-2015
New Jersey	✓	✓	2013-2014
New Mexico	✓		2013-2014
New York	✓	✓	2012-2013
North Carolina	✓	✓	2012-2013
North Dakota			2015-2016
Ohio	✓	✓	2013-2014
Oklahoma	✓		2015-2016
Oregon	✓		2014-2015
Pennsylvania	✓	✓	2013-2014
Rhode Island	✓	✓	2012-2013
South Carolina	✓		2014-2015
South Dakota	✓		2014-2015
Tennessee	✓	✓	2011-2012
Texas	✓		2016-2017
Vermont			2014-2015
Utah	✓		2013-2014
Virginia	✓		2012-2013
Washington	✓		2015-2016
West Virginia	✓		2013-2014
Wisconsin	✓		2014-2015
Wyoming	Under review		2016-2017

[a]From U.S. Department of Education (2015b). ESEA Flexibility.
[b]From U.S. Department of Education (2015c). Race to the Top fund.
[c]From National Council on Teacher Quality (2014b); Doore, B., Jorgensen, B., Doore, S., Mason, C.A. (2013). Teacher Evaluation and Professional Growth Systems.
[d]Some ESEA Waiver states have been given additional time, beyond what is indicated here, to implement the student growth requirement (Klein, 2015).

Appendix B

Letter to School District Superintendent (Phase I)

Dear Dr./Ms./Mr. _____:

I am contacting you to invite your district to participate in a study about teacher evaluation. I am a Ph.D. candidate at the University of Minnesota, and I am conducting my dissertation study on local interpretations of teacher evaluation policy.

Participation in the study involves my conducting telephone interviews with four to six individuals who have been involved in discussions and decisions about your district's teacher evaluation plan. Each interview is approximately 30 to 60 minutes in length. Ideally, I would like to interview one district administrator, one building principal, and at least two teachers. If you choose to participate, I will donate $250 to the _____ Foundation.

As you know, states all over the U.S. have adopted new or revised teacher evaluation policies in the past few years, and local school districts have responded by implementing state model plans or developing their own systems. At this critical time in evolution of this major policy initiative, it is important to understand the work of local educators as they design policies to meet state requirements. This study attempts to view teacher evaluation from a local perspective, to hear from educators on the front lines about the meanings of the policies and the challenges and opportunities these policies bring to local schools. The study is being conducted in two states, your state and one other.

In each state, I will interview educators in three school districts:

1. A district that has chosen to adopt a model approved or recommended by the state.
2. A district that has designed its own teacher evaluation model.
3. A district that has developed a teacher evaluation model with the aid of [the University of Missouri] [the Chalkboard Project].

Since your district has chosen to _____, yours is one of the districts I am contacting to fill that category.

I ask you to share information about the study with the individuals in your district who have been involved discussions and decisions about teacher evaluation so that I may conduct telephone interviews with them. The information from the interviews will be kept strictly confidential. Neither your school district nor the educators I interview will be identified in my dissertation or in any reports or publications that arise from the research.

I plan to call you within the next week to discuss this study. I look forward to speaking with you.

Sincerely,

Jane Gilles
Ph.D. Candidate
University of Minnesota

Appendix C

Consent Information Sheet (Phase I)

You are invited to be in a research study of local interpretations of state teacher evaluation policy. You were selected as a possible participant because you are part of the group in your school district that has been involved in discussions and decisions about the local teacher evaluation plan. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

This study is being conducted by: Jane Gilles, Department of Organizational Leadership, Policy and Development, University of Minnesota.

Procedures:
If you agree to be in this study, we would ask you to do the following things: Participate in a telephone interview of approximately thirty to sixty minutes to discuss your experiences with and thoughts about your state and local teacher evaluation policies. With your permission, the interview will be audiotaped so that the researcher can concentrate on the conversation during the interview.

Confidentiality:
The records of this study will be kept private. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records will be stored securely and only researchers attached to this study will have access to the records. Only the researcher will have access to the audio recordings of the interviews. The audiotapes will be destroyed at the completion of the research project, which is expected to end in fall 2015.

Voluntary Nature of the Study:
Participation in this study is voluntary. Your decision whether or not to participate will not affect your current or future relations with the University of Minnesota. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

Contacts and Questions:
The researcher conducting this study is: Jane Gilles (with the support of her academic advisor, Karen Seashore, Ph.D.). You may ask any questions you have now. If you have questions later, you are encouraged to contact Jane at 6213 Linda Lane, Lino Lakes, MN 55014; (651) 784-9009; jfgilles@umn.edu. Karen Seashore can be reached at 310D Wulling Hall, 86 Pleasant St SE, Minneapolis, MN 55455; (612) 626-8971; klouis@umn.edu.

If you have any questions or concerns regarding this study and would like to talk to someone other than the researchers, you are encouraged to contact the Research Subjects' Advocate Line, D528 Mayo, 420 Delaware St. Southeast, Minneapolis, Minnesota 55455; (612) 625-1650.

Please keep a copy of this information for your records.

Appendix D

Phase I Interview Protocol

Introduction:
Thank you for taking the time to speak with me today about your state's teacher evaluation policy and local efforts to meet the policy's requirements. This interview is part of a study about how educators in local school districts interpret state teacher evaluation policies. As part of this study, I am interviewing educators and policy leaders in this state and one other. Your responses will be kept confidential. This means that any information reported from this study will not identify you, your colleagues, your school district, or your school.

If you have no objection I will record this interview so that I can concentrate on our conversation. If at any point in the conversation you would like me to stop recording, please let me know. Also if there is a question you do not wish to answer, please tell me and we will move ahead. I am the only person who will have access to the audio recording of this interview, and it will be destroyed at the completion of the study. Do you have any questions before we begin?

Opening:

1. Tell me about your role in the school district. (Are you a teacher, principal, district administrator, or school board member? Other?)
   If the person is employed as an educator:
   a. How long have you been in education?
   b. How long have you worked in this school district?
   c. In which building do you work?
   If the person is a teacher:
   d. What grade/subject do you teach?
2. What has been your role related to your district's work in teacher evaluation?
3. How long have you been engaged in this work?
4. What is the current status of your district's teacher evaluation plan?
   Prompts if necessary:
   a. Is the design of the local plan complete?
   b. Is the plan pending approval of any kind?
   c. Are you piloting the plan?

Part A:

1. In your own words, what are the goals of your state's teacher evaluation policy?
2. In your understanding, what prompted your state to adopt a new policy on teacher evaluation?
   Prompt if necessary:
   a. Can you explain the factors that led to a new teacher evaluation policy in your state?

3. As someone who has worked closely with the state policy, are there any conflicts between the goals of the policy or between the goals of the policy and the reasons for its adoption?
   Prompts if necessary:
   a. For example, does it seem that some goals run at cross-purposes?
   b. Is there any conflict between the political discussions about the policy and the policy itself?
4. [For Missouri interviews only]; Has Amendment Three been a distraction as educators in your district have been implementing the new teacher evaluation system?

Part B:

1. In your opinion, what are the most important aspects – or requirements – of the state policy?
2. Do you think your opinion aligns with that of others involved in the design of your local plan in _____ school district?
3. Within your group, how have you resolved issues of confusion or disagreement about the goals and requirements of the state policy?
4. What have been the most challenging aspects of designing your local teacher evaluation plan?
5. Looking forward, what do you feel are the most promising aspects of your local plan as it evolves?
6. What aspects of your local plan concern you?

Part C:

1. Over the past few years, how did you and your colleagues in _____ school district learn about your state's teacher evaluation policy?
2. What documents have you relied on most in understanding the requirements of the state policy?
   Prompt if necessary:
   a. The policy language itself, publications from the state legislature, publications from the state department of education, resources from other organizations . . .
   b. Please describe the documents you have relied on most for understanding the policy.
3. Have there been any individuals or groups external to your school district that you have relied on to get information and resources about the policy?
   If yes:
   a. Which individuals or groups have provided the best information and/or resources about the policy?
   b. Tell me about this information and these resources. Have they been helpful? If so, how?

c. In your opinion, has the information and resources form external sources affected how you and your colleagues view the state policy, its goals, and its requirements? Please explain.

Part D:

1. Please provide a brief overview of your local teacher evaluation plan as it stands today.
2. Do you believe your local plan addresses the goals of the state policy and the reasons for its adoption? Please explain.
3. Are all the required components of the state policy addressed in your local plan?
4. Are any of the state policy requirements emphasized more strongly than others? If so, why?
5. Does your local plan include components that are not required in the state policy? *If yes:*
    a. Please describe these additional components.
    b. Why were they included in your local plan?

Closing:

Is there anything you would like to tell me about your state's teacher evaluation policy, your local plan, or your work on this issue that hasn't come up in our conversation?

Thank you so much for taking the time to talk with me today.

Appendix E

Letter to State-Level Policy Actor (Phase II)

Dear Dr./Ms./Mr. _____:

I am a Ph.D. candidate at the University of Minnesota, and I am contacting you about a study on teacher evaluation policy; I would like to interview you as part of this study. As you know, states all over the U.S. have adopted new or
revised teacher evaluation policies in the past few years, and school districts have responded by adopting new systems. At this critical time in the evolution of this major policy initiative, it is important to understand the work of educators as they develop local responses to state requirements. This study, which is my dissertation study, explores local interpretations of teacher evaluation policies. I am conducting the study in your state and one other.

In the first phase of the study I interviewed educators in three school districts in your state, and I am now beginning the second phase of the study – interviewing state-level policy actors. I am contacting you because your organization was mentioned as a helpful resource on teacher evaluation by educators in the three districts in your state that are part of the study. Would you be willing to participate in a telephone interview of approximately 30 to 60 minutes? The interview will address the following topics:

- Background about the policy, its development, and its implementation in local school districts.
- Your thoughts about the policy's goals and requirements.
- The efforts of your organization to support local educators in implementing the teacher evaluation policy.

Might you have some time in the next couple of weeks for a telephone interview? I am open most days and evenings all week – weekends, too! If you are willing to participate, please let me know what times might work for you. Note: I have enclosed [attached] an information sheet about the study. Please review it and let me know if you have any questions.

I plan to call you within the next week to discuss this study with you and to find out if you are willing to be interviewed. I look forward to speaking with you.

Sincerely,

Jane Gilles
Ph.D. Candidate
University of Minnesota

Appendix F

Consent Information Sheet (Phase II)

You are invited to be in a research study of local interpretations of state teacher evaluation policy. You were selected as a possible participant because educators in your state identified you or your organization as a source of information on your state's teacher evaluation policy. We ask that you read this form and ask any questions you may have before agreeing to be in the study.

This study is being conducted by: Jane Gilles, Department of Organizational Leadership, Policy and Development, University of Minnesota.

Procedures:
If you agree to be in this study, we would ask you to do the following things: Participate in a telephone interview of approximately one hour to discuss your views on your state's teacher evaluation policy and its interpretation by local school districts.

Confidentiality:
The records of this study will be kept private. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records will be stored securely and only the researchers attached to this study will have access to the records and the audio recordings of the interviews. The audiotapes will be destroyed at the completion of the research project, which is expected to end in fall 2015.

Voluntary Nature of the Study:
Participation in this study is voluntary. Your decision whether or not to participate will not affect your current or future relations with the University of Minnesota. If you decide to participate, you are free to not answer any question or withdraw at any time without affecting those relationships.

Contacts and Questions:
The researcher conducting this study is: Jane Gilles (with the support of her academic adviser, Karen Seashore, Ph.D.). You may ask any questions you have now. If you have questions later, you are encouraged to contact Jan; at 6213 Linda Lane, Lino Lakes, MN 55014, (651) 784-9009; jfgilles@umn.edu. Karen Seashore can be reached at 310D Wulling Hall, 86 Pleasant St SE, Minneapolis, MN 55455; (612) 626-8971; klouis@umn.edu.

If you have any questions or concerns regarding this study and would like to talk to someone other than the researcher(s), you are encouraged to contact the Research Subjects' Advocate Line, D528 Mayo, 420 Delaware St. Southeast, Minneapolis, Minnesota 55455; (612) 625-1650.

Please keep a copy of this information for your records.

Appendix G

Phase II Interview Protocol

Introduction:
Thank you for taking the time to speak with me today about your state's teacher evaluation policy and the efforts of local school districts to meet the policy's requirements. This interview is part of a study about how educators interpret state teacher evaluation policies. I am conducting this study in your state and one other.

For the first phase of this study, I interviewed educators in three small school districts in your state, and in these interviews, your organization was cited as a reliable source for information and resources about the state policy. Now, in the second phase of the study, I am interviewing state-level stakeholders who have supported local school districts in implementing teacher evaluation requirements.

Your responses to my questions will be kept confidential. This means that you will not be identified in any information reported from this study.

If you have no objection I will record this interview so that I can concentrate on our conversation. If at any point in the conversation you would like me to stop recording, please let me know. Also if there is a question you do not wish to answer, please tell me and we will move ahead. I am the only person who will have access to the audio recording of this interview, and it will be destroyed at the completion of the study. Do you have any questions before we begin?

Opening:

5. Tell me about your work role and the organization for which you work.
   Additional questions:
   a. How long have you been in this position?
   b. What was your previous position or employer?
   c. Have you ever worked in the K-12 school system?
6. What has been your role related to your state's work in teacher evaluation?
7. How long have you been engaged in this work on teacher evaluation?

Part A: Policy background

5. Next I have several questions about the history of teacher evaluation policy and practice in Oregon:
   a. What was the status of teacher evaluation policy and practice prior to the passage of Senate Bill _____ in 20___?
   b. Who was involved in moving the teacher evaluation policy agenda forward, and when did those efforts begin?
   c. How much influence did federal policies have on the evolution of teacher evaluation policy in your state?

6. In your own words, what are the goals of your state's teacher evaluation policy?
7. As someone who has worked closely with the state policy, do you perceive any conflicts between the goals of the policy or the reasons for its adoption?
   Prompts if necessary:
   a. For example, do any of the policy's goals run at cross-purposes?
   b. Is there any conflict between the political discussions about the policy and the policy itself?
8. How would you describe the state's attempts to secure and renew the federal ESEA waiver? How has this played out?

Part B: Policy requirements

7. In your opinion, what are the most important requirements of your state's teacher evaluation policy?
8. Do you think your opinion aligns with that of local educators across the state?
9. What do you believe are the most difficult challenges local educators have faced developing local systems to meet the policy's requirements?
   a. Do you feel that designing and implementing teacher evaluation processes has been any more difficult on small districts, rural districts, or districts with a high percentage of students in poverty

Part C: Resources for districts and educators

4. Please describe the efforts your organization has undertaken to support districts and educators in their work on teacher evaluation – over the past five to eight years.
   a. What motivated your organization to become involved in this work?
   b. What has been your strategy for communicating about teacher evaluation?
   c. What are the most important messages that your organization has tried to communicate about teacher evaluation?
5. Districts in Oregon report going to a variety of sources for resources and guidance on teacher evaluation. My interviews have indicated that, in addition to your organization, the following groups have been helpful to local educators: [Read list.]
   Questions:
   a. Are any organizations missing from the list?
   b. Not including your own organization, which organization do you believe has been the ***best*** resource on teacher evaluation?
   Only for first phase II interviewee in each state:
   c. Based on the information I received during the first phase of the study, I plan to interview the following individuals: [Read list of names and corresponding organizations.]
   d. Is there anyone else you would recommend that I speak to from the organizations I have listed?
6. In your organization's work on teacher evaluation, have you collaborated with other state-level stakeholders and groups?

If yes:

a. Please describe this collaboration.
b. How successful was the collaboration?
c. Do you believe your collaboration was helpful to educators at the local level? If so, how?

Part D: Moving forward

6. Are you optimistic that local teacher evaluation efforts will effectively address the goals of Oregon's teacher evaluation policy?
7. Do you anticipate that state policymakers are likely to revisit teacher evaluation in the near future?

Closing:

Is there anything you would like to tell me about your state's teacher evaluation policy or your work on this issue that hasn't come up in our conversation?

Thank you so much for taking the time to talk with me today.

Appendix H

Missouri and Oregon Education Data in Relation to U.S.

			U.S.	Missouri	Oregon
State Education Characteristics					
Nonwhite students[a]			46.3%	24.2%	29.7%
English learners[a]			8.5%	2.8%	9.1%
Students eligible for free or reduced-price lunch[a]			50.3%	44.7%	51.6%
Per-pupil expenditures / Rank among U.S. states[b]			$10,700	$9,597 / 30	$9,543 / 32
Students enrolled per teacher / Rank among U.S. states[c]			15.9	13.5 / 37	21.5 / 3
State Education Outcomes					
National Assessment of Educational Progress: Percent at or above basic/proficient[d]	Math	Grade 4	82% / 41%	83% / 39%	81% / 40%
		Grade 8	73% / 34%	74% / 33%	73% / 34%
	Reading	Grade 4	67% / 34%	70% / 35%	66% / 33%
		Grade 8	77% / 34%	78% / 36%	79% / 37%
ACT Average composite score / Percent of graduates tested[e]			21.0 / 57%	21.8 / 76%	21.4 / 36%
4-year graduation rate[a]	All students		81%	86%	69%
	American Indian/Alaskan Native		70%	82%	52%
	Asian		89%	91%	81%
	Black		71%	72%	57%
	Hispanic		75%	81%	61%
	White		87%	89%	71%
State Education Ratings					
Quality Counts (Education Week)[f]: Grade / Rank among U.S. states			C	C- / 33	C- / 41
Kids Count education scale (Annie E. Casey Foundation)[g]: Rank among U.S. states				23	35

[a]From U.S. Department of Education, ED Data Express. Data are from 2012-2013.
[b]From U.S Census Bureau. (2015b). *Public education finances: 2013*.
[c]From National Education Association *Rankings of the States 2014 and Estimates of School Statistics 2015*.
[d]From National Assessment of Educational Progress, National Center for Education Statistics. Data are from 2013. Retrieved from http://www.nationsreportcard.gov/reading_math_2013/#/.
[e]From ACT, Inc., 2014 National and state scores: http://www.act.org/newsroom/data/2014/states.html. Data are from 2014.
[f]From Editorial Projects in Education (2015). *Education Week Quality Counts*: http://www.edweek.org/ew/qc/index.html?intc=main-topnav. Rankings represent three categories of indicators: chance for success, school finance, and K-12 achievement.
[g]From Annie E. Casey Foundation *Kids Count 2015 Data Book*.

Appendix I

Missouri and Oregon Teacher Evaluation Policies and Systems

	Missouri	Oregon
Original policy adopted	1983	1979
Current policy adopted	2010 (amendment to 1983 law)	2011 (amendment to 1979 law)
Key policy requirements[a, b]	• Districts to evaluate each teacher through a performance-based evaluation system. • Districts to develop teacher standards; DESE to develop model standards. • Districts to adopt the state model or to develop/adopt/adapt a model that meets policy requirements.	• State to adopt teacher standards to include multiple measures, evidence of student growth, and opportunities for professional growth and individual goal setting. • Administrators, teachers, and exclusive bargaining representative to collaborate on local system design. • Districts to align local system to state framework, which outlines expectations.
State model or framework[a, b]	Missouri model evaluation system • Based on Missouri's seven principles of effective evaluation: 1. Research-based performance targets; 2. Differentiated levels of performance; 3. Probationary periods for new teachers; 4. Student performance measures; 5. Meaningful feedback; 6. Training for evaluators; 7. Use of evaluation results to inform personnel decisions • Three professional frames: 1. Commitment; 2. Practice; 3. Impact • Growth guide built on Missouri model teaching standards and indicators • Summative evaluation scoring clearly linked to personnel decisions • State-developed crosswalks to the Marzano evaluation system and the research of John Hattie	Oregon framework for teacher and administrator evaluation and support systems • Based on Oregon Core Teaching Standards • Differentiated performance levels (4) • Multiple measures with three categories of evidence: 1. Professional practice; 2. Professional responsibilities; 3. Student learning and growth. • Evaluation and professional growth cycle to include self-reflection, goal-setting, observation/evidence collection, formative evaluation, summative evaluation • Summative matrix • Aligned professional learning • Four state-approved evaluation rubrics (Danielson, Marzano, Marshall, Salem-Keizer School District rubric)
Percent of score based on student growth measures[a, b]	No specific percentage required	20% (Embedded in summative matrix scoring protocol)

Continued on next page

	Missouri	Oregon
Direct support organization model or framework[c, d]	Network for Educator Effectiveness (NEE) System (University of Missouri): • Assessments of classroom performance and other artifacts • Self-assessment • Professional development plan • Student perception survey • Online professional development modules • Training for evaluators • Data management resources • Access to comparison data	Chalkboard's CLASS Project: • Local plans are designed and implemented with coaching and support from Chalkboard • Teacher evaluation is one of four CLASS pillars: 1. Meaningful performance evaluation; 2. New career paths; 3. Expanded compensation models; 4. Targeted professional development

[a]From Missouri Department of Elementary and Secondary Education (2015c). Educator Growth Toolbox.
[b]From Oregon Department of Education. (2014a). Oregon framework for teacher and administrator evaluation and support systems: Revised for 2014-2015.
[c]From Curators of the University of Missouri (2015a). Network for Educator Effectiveness: About NEE.
[d]From Chalkboard Project. (2015a). CLASS Project.

Appendix J

Missouri School District Data in Relation to State

	Community Characteristics			District Characteristics				
	Proximity to major city[a]	Median household income[b]	Crime rate[c]	Nonwhite students	English learners[d]	Students qualifying for free or reduced-price lunch	Per pupil expenditures	District Enrollment
Missouri		$47,380	1,884.39	26.7%	3.5%	50%	$10,127	
Byrne Creek	More than 200 miles	Lower	Lower	Above	Below	Above	Above	1,000-1,499
Laurent	Less than 25 miles	Higher	Lower	Below	Below	Below	Below	1,000-1,499
Flores Valley	More than 100 miles	Lower	Higher	Below	Below	Above	Below	2,000-2,499

	District Outcomes					District Educators				
	Met AMO target		State exams where more students scored proficient or above than state avg	4-year grad rate	Grads attending post-secondary	Courses taught by "Highly Qualified" tchrs	Tchrs and admins with adv degrees	Tchrs and admins avg yrs experience	Avg salary	
	ELA	Math							Tchrs	Admins
Missouri				87.3%	68%	96.9%	58.9%	12.3	$47,844	$87,195
Byrne Creek	No	No	Less than half	Above	Below	Similar	Above	Higher	Lower	Lower
Laurent	No	No	More than half	Above	Similar	Similar	Below	Lower	Lower	Lower
Flores Valley	No	No	More than half	Above	Below	Above	Above	Higher	Lower	Similar

Note. Unless otherwise noted, data are from 2013-2014, accessed at Missouri Department of Elementary and Secondary Education (2014). Missouri comprehensive data system.
[a]Major city is defined as one of Missouri's three largest cities: Kansas City, St. Louis, and Springfield.
[b]From U.S. Census Bureau. (2015a). *2009-2013 American Community Survey 5-Year Estimates.*
[c]From World Media Group, LLC. (2015). USA.com U.S. crime index state rank; 2013 FBI data on incidence of major crimes normalized to 100,000 individuals (personal correspondence, March 30, 2015).
[d]Data are from 2014-2015 (personal correspondence, August 27, 2015).

Appendix K

Oregon School District Data in Relation to State

	Community Characteristics			District Characteristics				
	Proximity to major city[a]	Median household income[b]	Crime rate[c]	Nonwhite students	English learners	Students qualifying for free or reduced-price lunch	Per pupil expenditures	District Enrollment
Oregon		$50,229	1,522.07	35.9%	10.24%	54%	$10,256	
Meyers Grove	More than 75 miles	Below	Similar	Below	Below	Above	Above	500-999
Davies	Less than 50 miles	Above	Below	Below	Below	Below	Below	1,000-1,499
Nilsen-ville	About 75 miles	Below	Similar	Similar	Above	Above	Above	500-2,499[e]

	District Outcomes									District Educators						
	Met AMO target						State exams where more students scored proficient or above than state avg	4-year grad rate	Grads atten-ding post-secon-dary	Courses taught by "Highly Qualified" tchrs	Advanced degrees		Avg yrs experience		Avg salary	
	Elem ELA	Elem Math	MS ELA	MS Math	HS ELA	HS Math					Tchrs	Admins with doctoral degrees[d]	Tchrs	Admins	Tchrs	Admins
Oregon								69%	54.7%	98.3%	73%	6.6%	12.8	18.7	$58,185	$100,981
Meyers Grove	Yes	No	Yes	Yes	Yes	No	About half	Above	Below	Above	Below	Above	Above	Above	Similar	Similar
Davies	Yes	No	Yes	No	No	No	About half	Above	Above	Above	Below	Below	Above	Similar	Below	Below
Nilsen-ville	No	No	No	No	No	No	Less than half	Above	Similar	Above	Below	Below	Similar	Similar	Below	Similar

Note. Unless otherwise noted, data are from 2013-2014, accessed at Oregon Department of Education (2014b). Report card.
[a]Major city is defined as one of Oregon's three largest cities: Portland, Eugene and Salem.
[b]From U.S. Census Bureau. (2015a). *2009-2013 American Community Survey 5-Year Estimates.*
[c]From World Media Group, LLC. (2015). USA.com U.S. crime index state rank; 2013 FBI data on incidence of major crimes normalized to 100,000 individuals (personal correspondence, March 30, 2015).
[d]A Master's degree is required of all school administrators in Oregon.
[e]The enrollment range for this school district was broadened to protect the anonymity of the district.