Exploring Factors that Influence Teacher Perceptions of New Evaluation Policies

BY

JESSICA J. GOTTLIEB
B.S., University of Southern California, 2006
M.A., Loyola Marymount University, 2008

THESIS

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Policy Studies in Urban Education in the Graduate College of the University of Illinois at

Chicago, 2014

Chicago, Illinois

Defense Committee:

        Benjamin Superfine, Chair and Advisor
        Mark Smylie, Educational Policy Studies
        David Mayrowetz, Educational Policy Studies
        Pamela Quiroz, Educational Policy Studies
        Yue Yin, Educational Psychology

This dissertation is dedicated to my husband, Joe, for his unfailing love and support.

ACKNOWLEDGEMENTS

Over the past four years I have been fortunate to work with a number of enthusiastic, supportive colleagues. This dissertation would not have been possible with the support of my committee members, Dr. Ben Superfine, Dr. Mark Smylie, Dr. Pamela Quiroz, Dr. Yue Yin, and Dr. David Mayrowetz. I would like to thank Dr. Ben Superfine for being a wonderful advisor and chair, and a constant source of feedback, support, and advice. I would also like to thank Dr. Mark Smylie for his wisdom and kindness. Dr. Chris Miller has encouraged me to grow as a scholar and a teacher, and generously shared his office for the last four years.

I would also like to acknowledge the district leaders and teachers of River District. Thank you for your support and participation, without you this work would not have been possible.

JJG

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This study examines the factors that influence teacher perceptions of new teacher evaluation policies, as well as the relationships between the identified factors and overall perceptions. A mixed methods approach was used to provide both a broad comparison of different factors influencing teacher perception and an in-depth look at how individual teachers formulate their perceptions. A survey was used to collect data on the factors that influence teacher perception. To complement the data collected through the survey, semi-structured individual interviews were also used. The use of interview data provided deeper insight into the relationships between factors identified in the survey and overall teacher perceptions.

The extent to which teachers see an evaluation system as supporting the professionalism of the teacher workforce was the strongest significant predictor of overall teacher perceptions. However, while teachers viewed the local evaluation system as supportive of teacher professionalism, they felt the new state requirements pushed the system in a more bureaucratic direction and undermined the goals of the local evaluation system. Teachers no longer viewed the evaluation process as an opportunity to receive feedback and continue to develop as a professional. Rather, they now saw the evaluative process as something to be endured in order to stay employed.

The teacher's relationship with the evaluator and the teacher's perceived familiarity with the evaluation system were also significant predictors. Teachers who felt they shared with their evaluator an understanding of the purpose of evaluation were more likely to have a favorable opinion of the overall evaluation system. Likewise, teachers who felt confident and comfortable with the design of the evaluation system were more likely to have a positive overall perception of that system.

I. INTRODUCTION

A.   **Statement of the Problem**

Teacher evaluation policy is becoming an increasingly important part of education reform, largely because policymakers see teacher evaluation policy as a way to raise student achievement by improving the quality of the teaching workforce. Many aspects of schools, such as curriculum, class size, or even the physical arrangement of the classroom, have been tested and manipulated in an attempt to improve student learning and increase student achievement. Often, these reforms led to only mixed or limited results. Given these disappointing results, policymakers have continued to look for more effective ways of increasing student achievement. Policymakers have attempted to use teachers as a lever for raising student achievement in a variety of ways, including through different types of teacher workforce policy. Some policymakers advocate for more teachers and smaller classes. Others argue that the requirements for entering the teaching profession should be raised, in hopes of ensuring that only the best possible teachers are allowed into the classroom. Still others have argued that teacher education programs must be changed, or eliminated entirely. Education researchers have also started to explore these issues, largely by exploring the relationship between teacher quality and student achievement. A growing body of research points to the teacher as the key factor within the school for raising student achievement (Darling-Hammond, 2000; Goe, 2007; Konstantopoulos & Chung, 2011).

While it is possible that these new teacher evaluations policies may lead to the desired result – improved student learning – the history of education reform does not provide cause for optimism. In particular, the theory of action underpinning these new evaluation policies rests upon a faulty premise. These policies assume that the existence of higher stakes evaluations will

be enough to motivate teachers to change and improve their practice, and that teachers have the capacity to change. Yet, these new policies fail to take into account the individual motivations and characteristics of teachers as well as the organizational context in which teachers work. Contrary to the theory of action behind these policies, research on teacher motivations (Evans, 1996) has found that teacher motivations vary depending the characteristics of the individual teacher such as age. Additionally, previous work on teacher evaluation policies showed that teacher perception and response to evaluation policies varies with how the individual teacher views his or her school leadership (Stiggins & Duke, 1988). For example, trust, or the lack thereof, in school leadership could influence the effectiveness of a new teacher evaluation policy. Rather than being motivated to improve when faced with a new evaluation system, teachers might simply try to meet the minimum requirements or minimize their own struggles rather than risk a negative evaluation that could impact their job. Finally, the growth of alternative certification programs and other pathways to teaching may lead to a subset of teachers that might have different motivations or expectations for their careers.

Accordingly, the topic of this research study is how teachers view new teacher evaluation policies, in light of their individual characteristics, school contexts, and the policy design. This study was conducted in this particular district, which will be referred to as River District, because the changes made to their evaluation system are representative of the desired changes to teacher evaluation at the federal and state level. River District, a large suburban district in Illinois, has modified their existing teacher evaluation system [TES], to comply with recent teacher evaluation legislation in Illinois that is designed to improve the quality of the teacher workforce. Additionally, River District is a large suburban district and provides an opportunity to study the new evaluation system in the type of district that is not studied as often

as large urban districts, though it shares many student characteristics with those districts. Finally, the district and local teachers union collaborated on the initial development of the teacher evaluation system and the changes made in order to comply with the Performance Evaluation Review Act [PERA] and Senate Bill 7 [S.B. 7], so conducting the study in this district also provided an opportunity to consider how collaboration between district officials and the teachers union could influence the perceptions of new teacher evaluation policies.

Given all of these items, this research topic poses a problem because teachers are going to have a range of responses to evaluation policies, given that they have different individual characteristics, motivations, group identities, and organizational contexts. New research is needed to better understand the links between these elements and teacher perceptions of evaluation policies, as well as deeper insight into how teachers may respond to evaluation policies.

B.   **Purpose of the Study**
The purpose of this research is to identify the factors that influence how teachers formulate their perceptions of new teacher evaluation policies. This study will show the magnitude and direction of the correlation between different factors and teacher perception of new evaluation systems. It will also illuminate how and why various factors are related to teacher perceptions of evaluation policies. Accordingly, the study will be designed to address both the relationships between factors and perception as well as how and why those factors are related to teacher perceptions. Causal claims will not be possible because an experimental design is not possible for this particular topic at this time.

To address both areas of interest, a mixed-methods approach was used. A sequential explanatory design was used, which is characterized by two consecutive phases of research. The

first phase was a survey designed to measure the relationships between different elements of the policy, school, and teachers and teacher perceptions of new evaluation policies. The second phase was qualitative, with data collected through semi-structured interviews. This allowed the researcher to collect data on the thought process underlying teacher perceptions of the evaluation policy. Participants were selected based on the results of the first phase of the research. The specific research questions are as follows:

1. What are teachers' perceptions of new evaluation policies?

2. What are the factors that influence teacher perception of accountability-oriented evaluation policies?

3. What are the relationships between the identified factors and overall teacher perceptions?

C.   **Theoretical Perspectives**

Two theoretical perspectives serve as the theoretical foundation for this study. First, Spillane, Reiser, and Reimer's (2002) cognitive implementation framework provides a theoretical approach to understanding how teachers make sense of and perceive new policies in order to understand the implementation process. Second, Ingram, Schneider, and deLeon's (2007) model of the relationship between social construction and policy design illustrates how the way in which the construction of the target group of a particular policy informs many of the design elements of that policy, particularly the choice of policy instrument and distribution of benefits or burdens. Together, these perspectives combine to form a top-down and bottom-up model. This model shows both how policymakers' constructions shape the policy design, and how that policy design, combined with the organizational context and individual characteristics,

shapes the sensemaking process of the implementers. This, in turn, guides the implementation

and eventual success of the policy.

Spillane, Reiser, and Reimer's (2002) cognitive implementation framework argues that

teachers make sense of new policies based on their individual understandings and schemata, and

also within a social and organizational context. Furthermore, teachers respond to these policies

based on the signals the policy sends with respect to the problem that needs to be addressed,

what changes are required, and how these changes may address the problem at hand. These

different dimensions can and do interact with each other throughout the sensemaking process.

On the individual level, Spillane, Reiser, and Reimer argue that the teacher's beliefs about

teaching and learning, as well as a desire to maintain a positive self-image, affect how teachers

respond to new policies. At the organizational level, the norms, structure, and culture of the

organization will all play a role in how teachers perceive new policies, especially since those

policies are not implemented in isolation, but within a school context. Furthermore, the structure

of the organization can support teachers' sensemaking and positive perception formation, or it

can hinder it, depending on the opportunities teachers have to interact with their peers and

develop a shared understanding around the policy. Finally, the policy itself, and the implicit

images and values conveyed by the design choices, can also influence teacher perception.

Policy signals stem largely from the social construction of the target group of the policy

(Ingram, Schneider, & deLeon, 2007). Social construction is defined as "world-shaping exercise"

(Ingram, et al., 2007, p. 95), in which policymakers define people, objects, and events "through a

variety of images, stereotypes, and assignment of values" (p. 95). Broadly, target populations

that are constructed in a positive way, such as veterans, will result in policy designs that are

based around incentives and distribution of benefits, while groups that are constructed in a

negative way, such as felons, will see policy designs that are oriented towards punishment and use mandates as the policy instrument of choice. As the target groups are constructed in a particular way, this typically leads to differences in policy design based on the construction.

Conversely, policy designs convey how policymakers constructed the target population. This implicit image of the target population can be observed, and can influence how that population perceives the new policy. Spillane, Reiser, and Reimer (2002) argue that the representations constructed by policymakers are an especially influential factor in how teachers perceive new policies, because these representations either challenge or confirm the positive self-image of most teachers. In turn, this may shape the teachers' response to that policy, which will then either reinforce or change how policymakers view that population. This creates a cyclical relationship, in which policymakers design policy around a particular construction of a group, who in turn react in part based on the implicit image conveyed by the policy. Spillane, Reiser, and Reimer describe this process in another way:

> What is paramount is not simply *that* implementing agents choose to respond to policy but also *what* they understand themselves to be responding to. The "what" of policy only begins with policy texts, such as directions, goals, and regulations. (p. 393)

Combined, these theoretical perspectives point towards the factors that will influence how teachers perceive new policies, and the relationship between how policymakers construct and image of teachers, which informs the policy design, and how teachers will respond to those policies.

D.   **Limitations**

The findings from this study are limited by several factors. First, the study was limited to one district in Illinois. Accordingly, the uniqueness of this context makes the results difficult to

replicate in another context. Second, since the survey was distributed through the teachers union's email distribution list, only teachers who shared their email address with the union were contacted about this study. This was approximately 85% of the total population of teachers in the district. In the quantitative portion of the study, there was the possibility of a non-response bias, meaning that there is some bias between those who complete the survey and those who do not, especially given the low response rate. Finally, given the low response rate, the results from the quantitative analysis have limited statistical power and generalizability.

E.   **<u>Significance</u>**

This research will create new knowledge that could inform policymakers as they craft future evaluation policy. It will do this by clarifying how the design of the policy and the organizational context influence how teachers engage with these policies. By understanding how teachers engage with these policies, based on the design of the policy, the organizational context, and the characteristics of the individual teacher, policymakers can adjust their designs so that the desired response is more likely. This research can also inform ongoing implementation of the existing policies by providing information for school leaders about the elements of their organizations that are positively or negatively related with how teachers view new evaluation policies. Additionally, this study will contribute to the existing, and often fragmented, literature on teacher evaluation. In particular, this study, by virtue of its mixed-methods design, provides an opportunity to examine both the quantitative relationships between various factors and teacher perceptions and the thought-process of teachers when faced with a new evaluation system. Other researchers will be interested in this research because it will provide new insight into how teachers respond to evaluation policy within a highly politicized, often combative, context.

Hopefully, the findings from this study will point towards directions for future research, such as comparative studies between different districts and longitudinal approaches as these policies begin to be implemented nationwide.

II. TEACHER EVALUATION POLICY

A.   **History of Teacher Evaluation Policy**

New education reform attempts emerge out of the failures of previous policies to attain

the desired results. By understanding the failures of past policy, the current iteration of education

reform efforts, teacher evaluation policy in this case, can be better understood because the

current framing of the problem, the target of the policy, and the choice of instrument stem

directly from previous policies. Ingram, Schneider, and deLeon (2007) argue that extant policies

are especially powerful in the creation of new policy. In the case of teacher evaluation policy,

starting in the late 2000s a number of states have passed laws aimed at improving the quality of

the teacher workforce by the development of new teacher evaluation systems. Between 2009 and

2013, 26 states passed some type of legislation aimed at creating new teacher evaluation systems.

To understand how teachers are constructed in this wave of teacher evaluation policies, as well

as how the problem is defined, an examination of previous reform efforts is in order.

Certainly, the evaluation of teachers is not a new idea. In fact, teachers have been

monitored for quality and compliance since the beginning of the twentieth century (McLaughlin

& Pfeiffer, 1988). Historically, teacher evaluation was a function of the local education agency

(LEA), though by the mid-1970s and early 1980s some states such as Connecticut, North

Carolina, and Washington began to develop and implement teacher evaluation policies in hopes

of raising student achievement. However, as the federal role and state level role in education has

expanded, federal and state level education reforms have started to target teachers, and

specifically teacher evaluations, as a key leverage point for improving student achievement. The

current wave of new teacher evaluation policies stems, at least in part, from a push from the

federal level. The trajectory of federal education policy, as well state response to that policy, has

been shaped by the ways in which the federal government used various policy instruments to incentivize, mandate, or encourage states and districts to implement its desired policies.

By considering the trajectory of federal and state level education policy, the impetus for this push for new teacher evaluation systems can be better understood. Through tracing the changing problem definitions education policy attempts to address, it can be seen why leveraging teacher evaluation systems as a means to improve student learning is viewed as the correct policy solution. Furthermore, an analysis of the changing policy instruments also shows why the focus of federal and state education policy has shifted away from incentives and mandates, and towards a focus on building capacity at the teacher and school level. Finally, a history of federal and state education policy illustrates why teachers, rather than students, funding, school organizations, or other previous target groups, are now the focus of policy. Broadly, reform efforts can be seen as three policy cycles, with the new push for teacher evaluation representing the start of a fourth cycle.

1.   **First Policy Cycle: Students and Funding**

The first problem federal and state education policy attempted to solve was the persistent low achievement of low-income students. The Coleman report, as well as other social science, highlighted the deleterious effects of poverty on student outcomes, and heightened the sense that this was something the government can, and should, address (McGuinn 2006; Stein, 2004). Historically, local educational authorities [LEA] had almost completely control over education policy, and the federal government had little or no involvement. That began to change in the late 1950s, as the federal government began to enforce school desegregation decisions, most notably *Brown v. Board of Education*. By 1965, Congress passed the Elementary and Secondary

Education Act [ESEA] as part of President Johnson's War on Poverty, which expanded the federal government's role in schools and has remained in place ever since. The cornerstone of this legislation was Title I. This portion of the legislation granted funds to states, which in turn passed that funding on to LEAs, depending on the number of low-income students. The need to channel federal funds to LEAs drove the development of state education agencies [SEA]. The target group of this policy was the schools, which were thought to unable to effectively educate low-income students due to lack of funds.

Implicit in this policy are two assumptions: (1) low-income students have special education needs and (2) additional funding is sufficient to address those special educational needs. From the assumptions, problem – special education needs of low-income children are not being met – and the solution – more funding – emerges. It is clear that the solution is framed as one of increasing the inputs to the schooling process. Initially, there was little monitoring for compliance, and the federal government essentially took it on good faith that funds would be used for the desired purposes. Since this was an incentive, states could opt not to participate, and there were no consequences beyond missing out on Title I funds. As one might expect, the fidelity of the district level implementation varied widely, since there was little or no monitoring or enforcement at the state or federal level.

This focus on inputs, with a policy oriented around incentives to improve the education of low-income children lasted roughly until the publication of *A Nation at Risk*. This report sparked a shift away from a narrow focus on the special educational needs of low-income students by creating the sense of a national crisis. What was particularly striking about the report was that it argued that the education system was hurting all students, not just low-income students, by failing to provide an excellent education for all. Cohen and Moffitt (2009) note that

the view that student deficiencies and lack of resources were the root of the low achievement of students in poverty "was being replaced by the idea that the key deficit lay not with students nor with palpable resources but with schools that offered a thin diet of remedial education, rather than ambitious instruction" (p. 101).

2.   **Second Policy Cycle: Schools and Standards**

A Nation at Risk drew attention to the failure of the Title I funds to increase student achievement. In response to that failure and the rhetoric of *A Nation at Risk*, the next iteration of education policy reframed the problem as one of a lack of a systematic approach to student learning. Goals 2000, and later the Improving America's Schools Act [IASA], which was the 1994 reauthorization of the ESEA, were developed using the idea of systemic reform, a concept put forward by Marshall S. Smith and Jennifer O'Day (Smith & O'Day, 1990). Systemic reform framed the problem facing education as one of fragmented, incoherent policies that failed to create meaningful change in classroom practice (Smith & O'Day, 1990). The logic behind Goals 2000 was simple – student achievement failed to improve, despite the Title I funds, because the school system lacked a coherent, consistent approach to education. Like the ESEA, Goals 2000 used an incentive as its policy instrument. The federal government would give states funds to adopt the principles of systemic reform, which included the development of state content standards, performance standards, and curricula and assessments that were aligned to those standards. Success would now be measured not in terms of dollar sent to schools, but in terms of student achievement.

This emphasis on student outputs can be seen as the culmination of the gradual shift away from the focus on student inputs. In its initial form, Title I was designed to provide additional

funds so that schools could meet the unique educational needs of low-income students. The focus was on making sure that schools received sufficient inputs in terms of funding and resources. By Goals 2000, the focus had shifted way from providing the correct inputs to ensuring the right outputs – increased student achievements. Once again, a new problem definition – lack of coherent standards, curriculum, and assessments – and a new emphasis on student outcomes would require a new policy solution.

The education governance landscape shifted in response to these polices. As the federal government took a more active role and developed policies that moved towards the core work of schools, it encroached on territory that traditionally belonged to the states. The federal government also began taking a more aggressive approach to policy implementation and compliance. Likewise, state education agencies began taking an even more active role in setting local policy, since Goals 2000 and the IASA tasked them with developing standards and aligned assessments. Both levels of government borrowed strength from the other to advance their policy agendas (Manna, 2006). Moreover, both the federal government and state governments became more powerful, highlighting the fact that education governance is not a zero-sum game (Manna, 2006). Additionally, districts faced new pressures to comply with both state and federal reform efforts.

3. **Third Policy Cycle: Accountability**

The next major policy cycle, at all levels of government, was the No Child Left Behind Act (NCLB), which again, grew directly out of the failures of Goals 2000 and the IASA. The state response to Goals 2000 and the IASA varied, with some states embracing the reforms while others lacked the will or capacity to implement the desired reforms (McDonnell, 2005). Many

states still lacked rigorous curricula, high standards, and aligned assessments. It was assumed that states could implement the desired reforms, if they were only held accountable to do so. The problem was not that states could not develop aligned standards and assessments, it was that they would only do so if held accountable for the results. This logic extended to schools as well. Policymakers and officials believed that schools could provide excellent education for all students, if held accountable for student outcomes. The emphasis on student outcomes also completed the evolution away from policies that focused on inputs to policies that focused solely on outputs (McGuinn, 2006).

NCLB continued to use the same policy instrument at the IASA by requiring states to comply with the expanding requirements of the law in order to receive Title I funding. However, unlike the IASA, the expectations for compliance strongly enforced and for all purposes, the policy instrument used by NCLB can be considered a mandate. States would be allowed to design their own standards, develop their own assessments, and set their own proficiency benchmarks. However, under NCLB for the first time all states would be required to test students in grades three through eight. The results of these assessments needed to be disaggregated into student subgroups, such as racial and ethnic minorities, low-income students, and students with disabilities, and schools would be held accountable for the results.

Policymakers also recognized the role the teacher in meeting high standards and increasing student achievement. Venturing for the first time into workforce policy, NCLB also required that by the 2005-2006 school year, core subject areas would be taught by a "highly qualified" teacher. A highly-qualified met the following criteria: (1) fully certified by the state, (2) possessed a bachelor's degree, and (3) demonstrated subject matter competence as determined by the state (McGuinn, 2006; NCLB, 2002). Considering that by 2009 there would

be 7.2 million K-12 teachers in the United States, this was no small task (U.S. Census Bureau, 2011). Like the standards and assessments, states were allowed to determine their own requirements for subject matter competence.

Unfortunately, even for all of its explicit detail and threat of sanctions, or perhaps because of this, the implementation of NCLB proved to be challenging. One source of implementation issues can be found in the use of mandates as a policy instrument. Mandates impose costs both on the enforcer and on the object of enforcement (McDonnell & Elmore, 1987). The enforcing agency incurs costs because it must develop and maintain the infrastructure and human capacity to monitor compliance, as well as reacting swiftly and forcefully if the object of enforcement is out of compliance. The object of the enforcement, the states in this case, incur costs because they also must develop infrastructure and human capacity to carry out the regulations they have been mandated to follow. For example, under NCLB state education agencies, crucial actors in the implementation of the law, are forced to take on new roles including the development of statewide testing and accountability system (Mintrop & Sunderman, 2009), drastically expand the scope and responsibilities of the state education agencies.

While the use of a mandate can foster a culture of mere compliance (McDonnell & Elmore, 1987), the majority of NCLB's implementation issues stemmed from a paradox. NCLB required the schools and districts that were performing the worst to make the most drastic changes in their practice. Elmore (2002) succinctly sums up this paradox, "low-performing schools, and the people who work in them, don't know what to do. If they did, they would be doing it already" (p. 34). Not only did those schools lack the capacity to improve, they had already been the subjects of many previous reforms (Mintrop & Sunderman, 2009). This

realization led to a new understanding of the problem education reform policies needed to address. The problem was not that schools did not want to improve the education they provided, but rather it was that they lacked the skills and knowledge to do so. The issue was one of capacity, not one of will.

4. **Fourth Policy Cycle: Teacher Evaluation and Capacity**

The struggles implementing NCLB led to the re-framing of the problem as one of capacity. Schools, and in particular teachers, were failing to educate all students effectively not because they did not want to, but because they lacked the skills and capabilities to do so. Accordingly, the policy focus shifted away from forcing schools and states to do certain things to a focus on building the capacity of schools to improve student learning. Additionally, as it became clear that the strict mandate of NCLB caused many implementation issues, capacity building became the policy instrument for this next policy cycle. Once again, the choices made in the current cycle of policy design stem from the failures of the past cycles.

The political context during this cycle of policy design also supported the selection capacity-building as the instrument of choice. Due to political stalemate in Congress, the reauthorization of the ESEA seemed unlikely to happen quickly and the federal government was forced to find other ways to advance its education agenda, outside of the usual incentive provided by Title I funds. The American Recovery and Reinvestment Act was designed to provide stability to the U.S. economy after the 2008 financial meltdown and to stimulate the economy. It was a one-time appropriation, passed under extraordinary circumstances. Of the $787 billion total funds, $80 billion was set aside for education (ARRA, 2009). The State Fiscal Stabilization Fund (SFSF) received $53.6 billion of the education funding. To receive the funds,

states had to pledge to continue to make progress in areas outline by NCLB and the America Competes Act of 2007. These areas include standard and assessments, data systems to manage student achievement data, teacher effectiveness, and interventions for the lowest performing schools (U.S. Department of Education, 2009)

Despite the fact that it required these assurances, the SFSF's primary purpose was to prevent the reduction of the educator workforce, not to promote education reform. The primary lever for the new reform efforts was the Race to the Top (RTTT) competition. RTTT was a competitive grant program in which states competed for a share of $4.83 billion from the funds appropriated for the ARRA. To win the competition, states had to show that the state had "a comprehensive approach to education reform" (U.S. Department of Education, 2010b), which was the only absolute priority in the application. RTTT guidelines then identified five core areas to be addressed in the application, as well as an optional emphasis on science, mathematics, engineering, and technology education. While RTTT did not require any specific strategies or programs, the application, as well as the similarities amongst the winning applications, made clear the Obama administration's desired policies (McGuinn, 2010).

The first area to be addressed was State Success Factors, which required states to show that they had a political and legislative climate that was conducive to education reform. States demonstrated this by presenting evidence of stakeholder support, including the teachers unions and legislature, for the reforms proposed in the application, as well as evidence of past successes in education reform. The second area was Standards and Assessments, in which the state had to make assurances regarding the continual improvement of standards and assessments. The next area was Data Systems to Improve Instruction, which required states to lay out a plan to develop longitudinal data systems that allowed students and teachers to be linked. The area that was

17

worth the most points (138 out of a possible 500) was Great Teachers and Leaders, which required states to develop plans for new teacher evaluation systems that incorporated measures of student achievement into the way in which teachers were evaluated. The final section was Turning Around Lowest-Performing School, in which states had to explain their strategies for helping persistently low-achieving schools.

The weight given to the Great Teachers and Leaders section reflects the importance the Obama administration placed on increasing the effectiveness of teachers and school leaders. This echoes Elmore's (2002) argument that if teacher and school leaders knew how to raise student achievement, they would already be doing it. Secretary of Education, Arne Duncan, in a 2009 address to the National Education Association called the current teacher evaluation systems broken, and argued for a system that rewarded great teachers, supported struggling teachers, and addressed failing teachers (U.S. Department of Education, 2009b). The results of the competition illustrated this. Every winning state made assurances that it would support alternative pathways to teaching, incorporate student achievement data into teacher evaluations, and make decisions about tenure, promotion, and dismissal based upon teacher evaluations (Superfine, Gottlieb, & Smylie, 2012).

This foray into teacher evaluation policies illustrates the ways in which the Obama administration navigated the political environment in pursuit of their desired policies, and marks another step in the expansion of both federal and state level roles in district education policies. By making RTTT an optional competition, the Obama administration could avoid accusations of federal overreach. Additionally, the competitive structure of the program created political cover for reforms at the state level (McGuinn, 2012). Governors and lawmakers could argue that certain laws had to be passed if the state hoped to win a share of the RTTT money. Many states,

even states that did not end up winning the competition, passed new legislation regarding teacher evaluation policies, charter schools, and statewide data systems. Currently, more than 20 states have passed new teacher evaluation laws that incorporate measures of student growth into the evaluations, including Illinois(Mead, Rotherham, & Brown, 2012).

RTTT extended the reach of education reform policies in new ways. The emphasis on teacher evaluation policy was an unprecedented involvement with the workforce of the schools. While it did not mandate any particular policy, the RTTT made it clear that the Obama administration wanted student achievement data to form a significant role in teacher evaluations, and that those evaluations should form the basis of any personnel decisions made at the school level (Superfine, et al., 2012). In this way, the federal government was instructing districts how they should evaluate, reward, and dismiss the district's own employees. McGuinn (2012) neatly sums up this change, stating:

> This is a major shift and the long-term impact of a Democratic president taking on the unions over teacher accountability and school reform may prove to be one of the most important political legacies of RTTT. (p.147)

Indeed, RTTT proved to be a bargain, as even states that did not win a grant still passed legislation mandating new teacher evaluation policies.

The problem state and federal education policy seeks to address has evolved away from one of inputs to meet the unique educational needs of low-income students to one of capacity in the teacher workforce, based on both the failure of past policies and a belief that effective teacher always translates into increased student achievement. Teacher evaluation policy is seen as the obvious solution to this problem, because it is seen as a way to simultaneously improve teachers already in the workforce and to remove teachers that cannot or will not improve. These evaluation systems operate according to the belief that teachers will be motivated by the

evaluations and personnel decisions tied to these evaluations to improve their practice. Furthermore, it is believed that these evaluation systems will either build capacity in teachers or remove teachers that cannot build the necessary capacity to improve student learning.

However, these policies fail to take into account the vital role teacher perceptions of these new policies play in how teachers engage with and ultimately respond to new policies. The history of federal education reform efforts clearly shows that neither the promise of incentives nor the threats of sanctions can make teachers improve their practice if they do not have the capacity to do so. Mere compliance is much closer to the norm than sincere engagement with the proposed reforms. If teacher evaluation policies are to truly improve student learning by building the capacity of the teacher workforce, then the way in which teachers see and engage with these policies must be considered during the implementation process. Without this consideration, it is unlikely that this set of policies will live up to the expectations of policymakers and will only serve to set the stage for the next cycle of federal education policy.

B.   **Teacher Evaluation in Illinois**

Many states have responded to the federal push for new teacher evaluation policies by passing new legislation that forces districts to adopt new evaluation policies. This research focuses on Illinois because it is an example of a state that recently passed new teacher evaluation legislation, namely the Performance Evaluation Reform Act [PERA] and Senate Bill 7. PERA required the development of new teacher and principal evaluations that incorporated student growth data as a "significant factor" in the evaluations (PERA, 2010), as well as a framework that is aligned with the Illinois Professional Teaching Standards. Given the results of the evaluation, a teacher will be classified into four categories: unsatisfactory, needs improvement,

proficient, and excellent (PERA, 2010). These requirements are closely aligned with the RTTT guidelines, which is unsurprising since Illinois has applied to each of the three rounds of the competition, finally winning $42.8 million during the third round.

PERA also created the Performance Evaluation Advisory Council, consisting of teachers, principals, and other stakeholders, to develop a model evaluation system. The model evaluation system stated that student growth measures constituted 50% of the teacher's evaluation. Districts did not have to use the model evaluation system, but if they chose to develop their own system, it had to meet the requirements of PERA. Districts choosing to develop their own systems were required to form Joint Committees, which consisted of district officials and teachers, or bargaining officials where appropriate. The Joint Committees has 180 days from their first meeting to negotiate an acceptable evaluation system. If they failed to do so, the model evaluation would have to be used.

In June 2011, Illinois Senate Bill 7 was signed into law. Negotiations on this bill began immediately after PERA passed. Senate Bill 7 attached new stakes to the teacher evaluations outlined by PERA. In particular, the law now allowed for teachers to be dismissed if they receive two unsatisfactory ratings in a seven-year period. Additionally, vacant positions must now be filled based on merit, rather than on seniority, unless all other factors are equal. The requirements for tenure were also changed. Now, tenure could only be granted after three consecutive excellent ratings or if two of the last three ratings within in four-year period were proficient or excellent.

C.   **Teacher Evaluation in River District**

This study will be located in River District. Of the students in River District, 54.8% are classified as low-income and 21.7% are classified as English Language Learners. The racial and ethnic breakdown of the district is: 50.0% Hispanic, 32.0% White, 8.3% Asian-American, 6.7% Black, 2.4% Multi-Race, and 0.50% Native American. The graduate rate is 80% and 97 languages are spoken in River District homes. The district is home to the River District Teachers Association, which is an affiliate of the National Education Association. River District has implemented all elements of PERA, with the exception of student growth scores. Those are not mandated by PERA's implementation calendar until 2016.

The existing teacher evaluation system in River District is built around the Charlotte Danielson Framework for Teaching. The evaluation system has been modified to include all elements required by PERA, with the exception of including student growth scores as part of the teacher evaluations. Currently, the evaluations are based around observations, conferences, professional development, and portfolios. The evaluation system is robust, with distinct evaluation cycles designated for pre-tenured and tenured teachers. Pre-tenured teachers have taught for less than four full years with River District. In years one and two, pre-tenured teachers have at least one informal observation, between two and three formal observations, and one formative foundational observation. Formative foundational observations occur at the beginning of the school year with the goal of developing a collaborative, collegial relationship between the teacher and his or her evaluator. Reflective conferences are mandatory following each observation. Portfolios containing lesson plans, classroom rules, classroom diagrams, one informal assessment, and one formal assessment are also required each year. Informal assessments can by anything the teacher uses to gauge student learning, such as journal logs,

student work, or concept maps. Formal assessments may be any test, state, local, or teacher-created, or a graphic organizer. Pre-tenured teachers in years three and four have a similar process. Their formative foundational observation is optional unless the teacher is in a new placement. Additionally, the portfolio of a teacher in year three or four must include a goal setting sheet, which is used only for prompting reflection and facilitating discussion.

Tenured staff have a slightly different process. Informal observations are optional and can be requested. Formal observations are required at least once every two years. A summative conference must take place in the same year as the formal observation, in addition to the planning conference and reflective conference that is considered part of the formal observation process. Summative conferences are when the rating for the teacher will be determined. Tenured staff are also required to complete a professional development component, which can include participating in school or district initiative, inquiry into classroom practice, earning National Board Certification, pursuing a Ph.D. or Ed.D., or requesting support due to an assignment change. Each professional development selection can last between one and three years, after which another component must be chosen.

When the Illinois legislature passed Senate Bill 7 in June 2011, it put the district in the difficult position of adapting their existing system to meet the new state requirements. One of the largest changes was the requirement that teacher evaluation scores be used in making personnel decisions, including dismissal and recall rights. Like the existing system, the state requirements dictated four performance levels. However, the district performance levels did not translate perfectly to the state performance levels, requiring the district to develop a complex process that used the local district scores to determine the state evaluation scores for each teacher. Each teacher now received two scores for each evaluation cycle, one that reflected the requirements of

the existing local teacher evaluation system and one that reflected the state requirements. Essentially, the implementation of Senate Bill 7 created a set of dual evaluation systems. The local teacher evaluation system continued to emphasize professional growth and development, while the state teacher evaluation system was used in personnel decisions.

To evaluate teacher practice, River District used the Danielson Framework for Teaching. The four domains of teacher practice are planning and preparation, the classroom environment, instruction, and professional responsibilities. Within each domain, components of practice are listed, as well as descriptions of what unsatisfactory, needs improvement, proficient, and excellent practice for that component. For example, establishing a culture of learning is a component of the classroom environment domain. At higher performance levels, these components "describe teaching practice that is active, consistent with curriculum standards, differentiated, inclusive, engages students, aims at developing a community of learners, and incorporates teacher reflection" (Milanowski, 2004, p. 35).

Teachers are assigned a rating after their summative conference. The ratings, from highest to lowest, are as follows: Distinguished, Proficient, Basic, and Unsatisfactory. This rating is considered a reflection of the teacher's work and professional practice. In order to comply with PERA requirements, River District also assigns a second rating to teachers. This rating is called the state rating and this is the rating that is used for personnel decisions such as dismissal or rehiring. The match between the local rating and the state rating is not exact, but can generally be described as follows:

1. An Unsatisfactory under the local system is always an Unsatisfactory under the state system

2. A Basic under the local system is a Needs Improvement under the state system if the teachers requires a written, prescriptive program to improve his or her practice

    a. Otherwise, a Basic under the local system is a Proficient under the state system

3. A Proficient under the local system is a Proficient under the state system if the observation rubric is a mix of Basic and Proficient in different domains.

4. A Proficient under the local system is an Excellent under the state system if the observation rubric is a mix of Proficient and Distinguished in different domains

5. A Distinguished under the local system is always an Excellent under the state system

In short, the current focus on teacher evaluation policy is understandable, given the trajectory of state and federal reform efforts. However, as the focus of the reforms move closer to the work of teaching and learning, it is likely that familiar implementation challenges will arise. These reforms are no longer a matter of writing larger checks or asking states and districts to draft new standards or assessments. Instead, these reforms seek to improve the quality of the teacher workforce through a more meaningful evaluation process, while attaching new stakes to those evaluations. However, unless the ways in which teachers engage with these reforms is understood and taken into account, history makes it clear that it is likely that many of the same implementation issues, such as surface compliance or uneven fidelity to the policy, will arise.

III. CONCEPTUAL FRAMEWORK AND RELATED LITERATURE

A.   **Overview of Theoretical Framework**

New teacher evaluation policies require districts, school administrators, and teachers to implement new ways of assessing teachers, providing feedback to teachers, and using the information from teacher evaluations in new or expanded ways. These new evaluation systems have potential benefits such as evaluations that are less cursory and more closely tied to the actual quality of teaching, or at least a theoretical model of effective teaching. Furthermore, these new evaluation systems are designed to increase opportunities for teachers to receive feedback and support, as well as opportunities for professional development and teacher learning, this the goal that this will lead to instructional improvement. However, the policies also face potential pitfalls, such as mere compliance with the stated objectives, evaluations that are invalid or unreliable, or the development of perverse incentives that are out of line with the stated goal of these policies.

Literature on policy implementation provides some insight into the likelihood of these policies achieving their intended objectives – valid teacher evaluations used in meaningful ways to improve student achievement. Weatherly and Lipsky (1977) describe policy implementation as a process of adaptation and reinterpretation of the policy by the "street-level bureaucrats". Indeed, implementation has been described as a process of evolution, in which the policy is formulated and reformulated as objectives and resources change (Majone & Wildavsky, 1979). In their framework of the elements that influence policy implementation, Sabatier and Manzmanian (1980) point to such factors as the tractability of the problem, ability of the statute or policy to structure implementation, public support, and the attitudes of the constituent groups as important indicators to how successful a policy might be as it moves through different stages

of implementation. Honig (2006) highlights the key role of context and the interactions between people, places, and policies in policy implementation, arguing that to understanding implementation we must understand local context and its relationship to the demands the new policy places on implementers, not just the content of the policy or a universal set of factors.

While this research illuminates various factors that can shape policy implementation, it is primarily focused on the behavior of the implementers, and does not explore how the local actors develop the understandings, perceptions, and attitudes that play a large role in determining their eventual behavior. This research fails to take into consideration the sensemaking process local actors undergo when they are faced with new policy demands, and how this sensemaking on an individual and collective level may influence implementation. For example, Sabatier and Manzmanian (1980) state that attitudes of the constituent groups are important factors in implementation. By attitudes, they mean the positive perception of a new policy, and the extent to which the constituent groups support that policy. However, they do not address how those attitudes are developed, or how these attitudes might change in light of a particular policy. Likewise, Weatherly and Lipsky (1977) draw attention to the role of "street level bureaucrats" in shaping policy, but do not explore how these local implementers understand or perceive the policies they are implementing. In the case of teacher evaluation policy, teachers may engage with these policies in a variety of ways, not all of which will be congruent with policymakers' goals for the new teacher evaluation policies.

What is lacking is an understanding of how teachers develop their understandings and perceptions of a given policy, which is to say their sensemaking around that policy. To address this gap, Spillane, Reiser, and Reimer (2002) developed a cognitive framework for implementation as a support for more conventional ways of understanding implementation,

which largely focuses on the actions of implementers and not their beliefs or understandings. They argue that focusing on behaviors alone is insufficient to understanding policy implementation. Rather, the way in which local actors perceive and understand the proposed change is vital to explaining the implementation process. They note when it comes to policies, "what we see is influenced by what we expect to see" (Spillane, Reiser, & Reimer, p. 395). In their study on teacher evaluation policy in Finland, Tuytens and Devos (2010) apply this framework to understand how school leadership influences teacher perception of evaluation policy.

This framework posits that what a policy means for an individual can be understood through the interaction of three different levels of implementation: the individual's own sensemaking, the situation or organizational context that influence the individual's sensemaking, and the representations and policy signals of which the individual attempts to make sense (Spillane, Reiser, & Reimer, 2010). At each level, individual, organizational, and policy, there are a variety of factors that influence this sensemaking and development of perceptions. This framework presents a bottom-up perspective of policy implementation.

This sensemaking process can thought of as one in which individuals interpret policy signals, such as the design of the policy, choice of policy instrument, and definition of the problem, both through their own understandings and within their social context. Therefore, certain elements of the policy itself play a key role in the eventual perception of the policy, even though teachers make sense of a policy within an organizational and social context. Ingram, Schneider, and deLeon (2007), in their description of the relationship between social construction and policy design, argue that the way in which the target group of a policy is defined or constructed by policymakers informs and influences the eventual design of a

28

particular policy. They define social construction as "a world-shaping exercise" and the "ways in which the "realities" of the world are defined" (Ingram, et al., 2007, p. 95). Through this process, narratives about problem definition, target groups, and solutions are developed (Stone, 2001). In the case of teacher evaluation policy, these narratives convey, at least in part, what policymakers believe about teacher and teaching, and the images of teachers that the policy is built upon. In short, what the policymakers believe about the group at hand, in this case, teachers, will directly influence how they design a policy, and the design of a policy can reveal how policymakers constructed teachers.

The construction by policymakers of the target group, or policy image of a group, directly influences the choice of policy instrument, allocation of burdens or benefits, and opportunities for participation. Policy instruments are mechanisms used to translate the goals of a policy into real-world actions (McDonnell & Elmore, 1987) by encouraging desired actions through the use of positive or negative incentives. McDonnell and Elmore (1987) define the four policy instruments most commonly used in education policy as mandates, inducements, capacity-building, or system-changing. Mandates are rules generally intended to force compliance. Inducements are incentives, usually monetary, to for the target of the policy to act in the desired way. Capacity-building instruments attempt to elicit the desired actions by increasing the financial, social, or human capital of the implementer. System-changing instruments transfer "official authority among individuals and agencies in order to alter the system by which public goods and services are delivered" (McDonnell & Elmore, 1987, p. 134). The choice of a given policy instrument directly shapes the design of the policy, because that choice determines how the policy will attempt to influence the target population to act in the desired ways.

29

Policy tools are based upon a set of behavioral assumptions about the target population (Schneider & Ingram, 1990). Additionally, the choice to use a particular policy instrument reveals something about the policymaker's values and perception of the target group (Linder & Peters, 1989). This is precisely the type of policy signal that teachers will respond to through their sensemaking process, as described by Spillane, Reiser, and Reimer (2002). By considering social construction of target groups in light of a cognitive implementation framework, it becomes clear that it understand how teachers will perceive new evaluation polices requires an understanding of the image of teachers and teaching implicit in the policy at hand, because that image drove the policy design. In turn, the policy design conveys important signals to the teachers that they will have to make sense of, both individual and within their organizational context. The alignment, or lack thereof, between the policy image of teachers, the policy signals sent by the policy's design, and the teachers' own understandings will shape the implementation process for better or for worse. This top-down perspective can be combined with the bottom-up perspective presented by Spillane, Reiser, and Reimer (2002) to present a comprehensive view of the implementation process.

Overall, it is clear that teacher perceptions, that is, the understandings and attitudes towards the policy at hand, play an important role in the eventual implementation of a policy. In a sense, perception serves as a vehicle through which the intended form of the policy is translated into the actual policy at the local level. Teachers will act in response to their perceptions of these policies, so the extent to which that perception is positive or negative will have marked consequences for the eventual implementation and success of a particular policy. Furthermore, the image of teaching conveyed by the policy's design and its congruence with the teachers' own identity and understandings about the work of teaching Hence, it is important to

understand how teachers perceive these new evaluation polices, as well as the ways in which these perceptions are developed, in order to better understand the likelihood of these policies producing the desired response.

The cognitive implementation framework, as well as the literature on social construction of target groups, provides a base for understanding the factors that may influence how teachers perceive new teacher evaluation policies, and the process through which they develop their perceptions. The cognitive implementation framework indicates that three levels of factors – individual, organizational, and policy level – may influence how teacher perceive new policies. Furthermore, the literature on social construction of policy indicates that the alignment between the image of teaching implicit in these new teacher evaluation policies and the teachers' own understandings of teaching, as mediated through the organizational context, may also shape how teachers view these policies. The following sections will examine each level of factors, as well as how those factors may interact across different levels. Figure 1 provides an overview of the factors that will be considered, based on an examination of the literature. By considering these factors, a fuller picture of how teachers may perceive new evaluation policies will emerge.
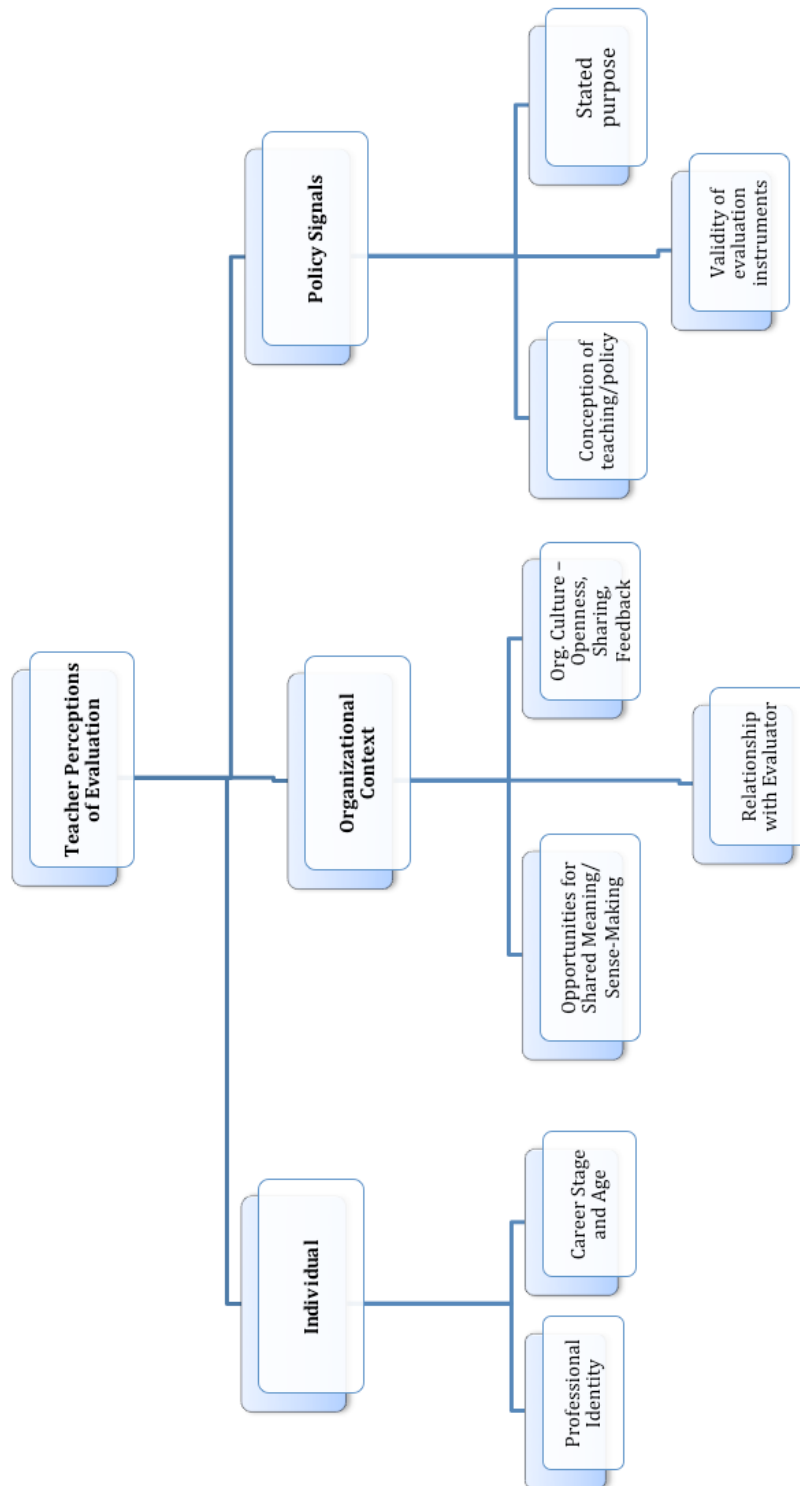
Figure 1. Factors that may influence teacher perception of new teacher evaluation policies

B.   **Policy Signals**

Policy signals from the new teacher evaluation policy may influence how teachers perceive it, and thus how they respond to it. These elements include the purpose of the evaluation system, the type of standards used to evaluate teachers, and the conception of teaching and image of teachers embedded in the evaluation system. Additionally, elements such as validity and fairness may affect the perceived legitimacy of the teacher evaluation system. More broadly, these characteristics convey messages about the problem that needs to be addressed, who or what will be changed in order to address the problem, and the extent to which those involved will be required to change (Spillane, Reiser, & Reimer, 2002). As these changes become more fundamental, the extent to which implementers need to develop new understandings and restructure their schema increases (Spillane, Reiser, & Reimer). Accordingly, the more difficult the change or policy implementation may be.

1.   **Purpose of the Evaluation System**

The first element of the evaluation system that can influence teacher perception is the purpose of the evaluation system. Purpose is simply the espoused goal of the evaluation system. It should be noted that the perceived purpose and intended purpose may not always match, in which case, the perceived purpose becomes the one upon which a teacher will act. For example, in their study of a new teacher evaluation system in Cincinnati, Ohio, Milanowski & Heneman (2001) found a variety of perceived purposes for the evaluation system, ranging from eliminating weak teachers to improving student achievement. Broadly, the stated purpose of an evaluation system can be categorized as either accountability or improvement (Duke & Stiggins, 1990). An accountability-oriented evaluation system "involves the collection of data to determine the extent

to which teachers have achieved minimum acceptable levels of competence" (Duke & Stiggins, 1990, p 116). In contrast, evaluation systems with the purpose of improvement focus on assisting "teachers who are at least minimally competent in continuing to grow" (p. 117). The espoused purpose of the evaluation system can be found through a variety of policy texts, including the objectives, instruments, and expected uses of the evaluation system (Spillane, Reiser & Reimer, 2002).

The stated purpose of an evaluation system affects teacher perception for two reasons. First, a cognitive approach to implementation suggests that teachers must construct understandings of new policies and that these understandings must align with policymakers' goals if those goals are going to be enacted (Spillane, Reiser, & Reimer, 2002). Teachers might believe that the purpose of an evaluation system should be improvement, rather than accountability, or some combination of both (Kimball, 2002), and these beliefs might vary in their alignment with the stated goal of the policy. Second, the stated purpose will influence perceptions because that purpose will shape the eventual use of the evaluations. It is likely that teachers will have a different view of an evaluation system that they know will be used to make personnel decisions, instead of just being used for feedback and improvement. Teachers will perceive differences in expectations and in the standards used to evaluate teachers. The stakes attached to the evaluation will also be different.

One of the most common stakes attached to teacher evaluation systems oriented around accountability is performance-related pay or merit pay. Performance-related pay is a shift away from traditional teacher compensation structures, in which teachers are paid based on years of experience and educational level, according to a salary scale that is often collectively bargained. Since the 1980s there has been a growing interest in implementing performance-related pay in

K12 public schools. The continual resurrection of these ideas would seem to indicate that they have a track record of success, or are popular with teachers. Neither of these statements is true, and in fact, the incorporation of these ideas could negatively influence teacher perception of a new evaluation system.

Studies have shown that teachers do not support merit pay for a variety of reasons. One of the primary reasons merit pay programs have not been successful is that the programs violate teacher professional norms and values through the use of competitive evaluation standards (Peterson, 1990; Smylie & Smart, 1990). In their study, Smylie and Smart (1990) found that teachers were concerned about the effect of merit pay programs on workplace dynamics, since merit pay, by definition, differentiates amongst teachers, reducing the equality that teachers prize and fostering competitiveness. Indeed, Peterson notes, "it is not considered good form to go public with exemplary practice" (p. 107), and to do so is a violation of workforce norms.

Another stake attached to teacher evaluation systems built around accountability is the use of these evaluations for personnel decisions such as tenure or dismissal. Using evaluations in this way amplifies the importance of the evaluation, since a teacher's livelihood could be affected by a series of poor evaluations. Depending on the extent to which the teachers view the measures and instruments of the evaluation as valid, using evaluations in this way could affect how teachers view the evaluations. Now, instead of viewing the evaluation process as an opportunity for feedback and improvement, teachers may feel especially threatened by the process, since important career milestones hinge on the evaluations.

In practice, however, these two purposes often exist side-by-side in the same evaluation system, though one, usually accountability, is prioritized publically. Thus, it is likely that it is not so much the choice of the purpose that will influence teacher perception, but rather how school

administrators and evaluators manage the inherent tension between the two goals (McLaughlin, 1990). This dilemma can be seen in the RTTT guidelines, in which school districts are simultaneously exhorted to develop support systems for struggling teachers, as shown by their evaluations, and at the same time use those same evaluations to make personnel decisions such as dismissal. In addition, while policy texts such as guidelines and objectives convey the purpose of the evaluation system, teachers also receive information and messages about these systems from other groups such as teachers' union, the media, and their colleagues. Darling-Hammond, Wise, and Pease (1983) note, "this tension between evaluation goals is in part a reflection of the differences among evaluation constituencies. These stakeholders have divergent views on the primary purpose of teacher evaluation, and hence, of what constitutes a successful evaluation system" (p. 22). These messages, coupled with the ways in which the evaluators manage the tension between the divergent goals of the system, form the basis for how teachers develop their perceptions of the purpose of the evaluation system.

2.   **Evaluation Standards and Measures**

Not only does the stated purpose of the evaluation system influence teacher perception, the standards against which the teachers are judged also influence perception. The choice of standards reflects the policymakers' beliefs about how best to improve student learning, and are closely tied to the espoused purpose of the evaluation system (Bacharach, Conley, & Shedd, 1990). Standards for evaluation fall into one of three categories: minimum standards, competitive standards, and developmental standards. Minimum standards are exactly what they sound like – standards that set a minimum level of performance or competency (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1985). Competitive standards "provide incentive and rewards for

some form of superior performance or for particular kinds of performance" (Bacharach et al., 1990, p. 137). Rightly, teachers perceive these types of standards as fostering competition and perhaps undermining collegiality (Smylie & Smart, 1990). The final type of standards encourages a focus on teacher growth (Bacharach et al., 1990). Naturally, the use of this type of standard will shape teacher perception of the evaluation system, perhaps framing it as a system as something to be embraced, rather than feared.

Regardless of the type of standards chosen, the way in which they are communicated to teachers will also influence teacher perception. Duke and Stiggins (1990) argue that clarity of the performance standards, the extent to which the teacher was made aware of the standards, and the extent to which the teacher agrees with the standards are positively correlated to the perceived quality and impact of the evaluation experience. Peterson and Comeaux (1990), in their study of teacher evaluation systems, found that, "teachers' rating of different systems were related to teachers' beliefs about what constitutes good teaching" (p. 22). Additionally, the perceived teacher input or guidance in the development of standards or rubrics may also shape how teachers ultimately view these evaluation instruments.

3. **Conceptions of Teaching and Images of Teachers**

The purpose of evaluation, as well as the standards used in the evaluation process, stem from a deeper conception , or policy image, regarding the work of teaching. Wise et al. (1985) argue, "a teacher evaluation system must define the teaching task and provide a mechanism for judging the teacher" (p. 65). Different conceptions of teaching then necessarily lead to different ideas about what good teaching is, how good teaching can be identified, and how teaching can be improved (Darling-Hammond, Wise, & Pease, 1983; Peterson & Comeaux, 1990). The teaching

task can be defined in four ways: labor, craft, profession, and art (Darling-Hammond, et al., 1983; Wise et al., 1985), each with its own implications for evaluation. The conception of teaching as labor or craft imply a more rationalistic or mechanistic school organization, while the conception of teaching as profession or art imply a school organization that is more organic or less rationalistic (Darling-Hammond et al., 1983; Weick & McDaniel, 1989).

The conception of teaching as *labor* envisions the work of teaching as routinized and rationally planned. This definition conjures up a vision of factory work. Given this definition, the evaluation system will need to analyze the teacher's work product, in the form of lesson plans, classroom performance, and student achievement results (Wise et al., 1985). In this case, the evaluation system hopefully reveals specific effective practices, and encourages all teachers to follow these practices for the desired results. Viewing the work of teaching as a type of labor minimizes the need for specialized knowledge on the part of the teacher. While a focus on the teacher's work product, usually student test scores or identifiable classroom practices, implies that good teaching is a sufficient component of student performance, this focus also characterizes teaching as set of procedures and protocols that leads to a predictable outcome (Bacharach, et al., 1990). The conception of teaching as a *craft* goes beyond the conception of labor by acknowledging that teaching requires a set of specialized techniques (Wise et al., 1985). Here, an evaluation system would be oriented towards assessing the teacher's skills by making sure that they had all necessary techniques at their disposal.

If the conception of teaching includes the exercise of judgment as well a set of specialized techniques, then it becomes the conception of teaching as a *profession* (Wise et al., 1985). Achinstein and Ogawa (2006) define teaching professionals as those with:

> specialized expertise, who have discretion to employ repertoires of instructional strategies to meet the individual needs of diverse students, hold high expectations for

themselves and students, foster learning communities among students, and participate in self critical communities of practice. (p. 32)

Again, the type of evaluation system shifts. For a conception of teaching as a profession, the evaluation system should assess "the degree to which teachers solve professional problems competently" (p. 65). Additionally, under this conception of teaching, a teacher's peers would develop the evaluation standards, as with other professions such as medicine or law. The final conception of teaching is teaching as *art*. This conception builds upon the conception of teaching as a profession by incorporating personal insight, creativity, and improvisation into the use of judgment and specialized skills (Darling-Hammond et al., 1983; Wise et al., 1985). An evaluation system build upon this conception would incorporate both critical assessment from others as well as self-assessment (Wise et al., 1985).

The conception of teaching underpinning a policy is usually conveyed implicitly through aspects of the policy such as the goals, regulations, objectives, and requirements. Spillane, Reiser, and Reimer (2001) argue that these conceptions, or representations, are incomplete and subject to the interpretation of the individual teacher or implementer. They state, "the meaning of external representations exists fully only when individual decisions and actions are based on what has been actively interpreted and constructed as a result of interaction with the artifact" (p. 416). Therefore, conceptions of teaching are initially, and only partially conveyed, through the requirements, objectives, and goals of the policy itself, and then are expanded as the individual teachers draw upon their own understandings and experiences to develop what they believe the policy's conception of teaching to be. It is through this process that other groups, such as the teachers' union or advocacy organizations, play a role in shaping what teachers ultimately believe about the view of teaching implicit in a given policy.

Each conception of teaching would influence how teachers perceived an evaluation system, either by minimizing the work of teaching through a conception of teaching as labor, or by elevating it by viewing it as a profession or art. If most teachers view themselves as professionals (Sachs, 2001), an evaluation system built upon a conception of teaching as a labor or craft could certainly diminish the likelihood of teachers responding the evaluation system in a positive way. Additionally, each conception of teaching also has a different definition of success, ranging from the effective production of some output to the full use of one's creative and intellectual abilities. Certainly, the definition of success will also influence teacher perceptions of the evaluation system, especially if the system's definition of success does not align with the teacher's definition.

These differing conceptions also play a role in how a district or state approaches teacher evaluation and other types of education policy (Wise & Darling-Hammond, 1984). Each conception carries with it a set of assumptions about the work of teaching and how it is best evaluated. Broadly, conceptions of teaching as labor or craft are aligned with a bureaucratic approach to teacher evaluation, while conceptions of teaching as a profession or art are aligned with a professional approach to teacher evaluation. Tschannen-Moran (2009) argues that these two approaches to evaluation draw upon differing assumptions about the capacity of workers, in this case teachers. The bureaucratic approach to school reform "embodies an implicit distrust of teachers and the contributions they have to offer", while the professional approach is "grounded in trust" (Tschannen-Moran, 2009, p. 220). Adler and Borys (1996) similarly describe two approaches to organizational design, one that views the worker as the source of problems to be solved, while the other approach views the worker as a skillful and valuable source of solutions. Given these differing assumptions, these two approaches to evaluation, bureaucratic and

professional, will lead to evaluation policies that either promote or hinder positive teacher

perceptions of the policies.

A bureaucratic orientation to teacher evaluation is built around a conception of teaching

as industrial labor or craft. As such, a bureaucratic teacher evaluation system will focus on

quantifiable outputs, such as student test scores, and will emphasize close monitoring of

teachers' work and compliance with work rules and policies (Tschannen-Moran, 2009).

Bureaucratic teacher evaluation will also be generic, often utilize a checklist, and take a one-size-

fits-all approach to evaluation (Darling-Hammond, 1990). Evaluators will be administrators,

usually principals, not other teachers or content area experts (Wise & Darling-Hammond, 1984).

Therefore, evaluations in a bureaucratic system will be superficial, fail to provide helpful

feedback, and ensure only minimal competence (Wise & Darling-Hammond, 1984).

In contrast, a professional teacher evaluation system will provide detailed, actionable

feedback and will be "designed to meet teachers' needs for guidance in addressing specific

problems of classroom practice" (Wise & Darling-Hammond, 1984, p. 31). A professional

approach to teacher evaluation views teachers as experts who can contribute to the design of the

system, so teachers will be involved in the every step of the system, from design to conducting

the evaluations. A professional teacher evaluation system is built around shared professional

standards and is designed to support professional growth and development (Darling-Hammond,

1990). Darling-Hammond (1990) states:

> The most appropriate strategies for growth-oriented evaluation are personal goal setting
> and self-evaluation, joined with peer-mediated and situationally relevant reviews of
> practice, conducted by teachers sharing similar expertise and teaching assignments. (39)

This differs markedly from the bureaucratic approach to teacher evaluation. Accordingly, the conception of teaching and accompanying approach to evaluation may play a significant role in how teachers view new teacher evaluation policies.

In addition to the conception, or image, of teaching found in a policy, the image of teachers themselves presented by a given policy can also influence perception of that policy. This is distinct from the conception of the work of teaching, rather it is the idea of the teachers -- who they should be and what they should do – themselves that is implied by the policy. Jansen (2001) describes policy images as "official projections through various policy texts of what the ideal teacher looks like" (p. 242). These policy images provide, implicitly or explicitly, a way of understanding how policymakers view teachers when formulating policy. These images can range from idealized images of teachers to negative portrayals, and will influence the eventual design of the policy, such as the selection of policy instruments used in implementation (Ingram, Schneider, & Deleon, 2007). For example, in their study of NCLB Cochran-Smith and Lytle (2006) found NCLB to be premised upon a narrowly constructed policy image of teaching that prioritized subject matter knowledge and data-driven instructional decision making. This image of teaching can be directly linked to the testing and accountability structure of NCLB. Because policy solutions are designed around particular policy images, such as in the case of NCLB, the perception of the policy itself may be influenced by the perception of the policy image underpinning the policy.

C.   **Organizational Context**

The characteristics of the policy itself are not the only factors that influence how teachers may perceive new teacher evaluation systems. These evaluation systems do not exist in isolation.

Rather, teachers make sense of these new policies within the context of their organizations. These organizations share some characteristics that are common to schools, but also each have their own set of structures and culture, all of which will play a role in how teachers make sense of, and ultimately perceive these new evaluation policies. In their cognitive implementation framework, Spillane, Reiser, and Reimer (2001) argue that the situation or context of an individual, which in this case is the school, is crucial to understanding how he or she makes sense of new policies. Accordingly, this section articulates the key organizational factors that may influence how teachers make sense of and perceive new evaluation systems.

Implementation of new teacher evaluation systems is unlikely to be consistent across schools or within schools. Weick (1976) described schools as loosely coupled systems, in which different elements of the organization are responsive to one another, but also remain distinct. The structures within the school, such as grade level teams or academic departments may very well respond differently to new evaluation policies. Furthermore, Coburn (2001) argued that sensemaking can vary across informal groups as well as formal groups. She found that teachers build informal groups around those with similar worldviews, which can lead to variations in how different informal teachers respond to the same policy signals. The structure of the school can create opportunities for teachers to interact with those who interpret the policy in different ways, or it can limit those interactions and reduce the opportunity for teachers to be exposed to understandings and interpretations that are different than their own (Spillane, Reiser, & Reimer, 2001).

Ultimately, teachers are likely to be part of several formal and informal groups within their school, all of which influence the sensemaking about a particular policy. However, in the case of teacher evaluation policies, the literature points towards several key factors stemming

43

from formal organizational structures that may influence teacher perceptions and sensemaking of these policies. First, the greater the alignment between the teacher evaluation system and the goals of the organization in which the program is embedded, the more likely it is that the teacher evaluation system will be seen as effective and useful. Iwanicki and Rindone's (1995) description of the relationship between organization context and evaluation systems bears repeating:

> Our contention is that effective teacher evaluation programs are productive to the extent that they are consistent with and integrated into the organization context of the school system and its schools. There are certain beliefs about the teaching-learning process in schools that the leadership of the school system holds to be true. These beliefs help shape the organizational context of the system. There are certain beliefs about the teaching-learning process which teachers, supervisors, and administrators at the school building hold to be true. These beliefs help shape the organizational context of the school building. To the extent that these system and building level beliefs are consistent with the assumptions underlying more effective teacher evaluation practices, there is a higher probability that the teacher evaluation process will be more productive. (p. 78).

Thus, the members of the organization will perceive the new evaluation system through the organizational context of the school, especially the beliefs and culture of the organization.

1.   **<u>Organizational Culture – Openness to Feedback and Sharing</u>**

As Iwanicki and Rindone (1995) illustrate, the fit between the beliefs of the school organization and the underlying assumptions, such as purpose and conception of teaching, in the evaluation system will largely determine both teacher perceptions of the evaluation system, as well as its ultimate effectiveness. However, beliefs about teaching and learning vary across schools, as well as within schools. Indeed, every organization has its own culture, which is simply the "pattern of shared basic assumptions the group learned as it solved its problems of external adaption and internal integration" (Schein, 1996, p. 12). Certainly, teachers constitute a part of each school organization, and therefore will be influenced by the assumptions and culture

44

of their organization (Firestone and Louis, 1999; Schein, 1996). While these assumptions may vary from organization to organization, there a certain patterns that can be identified in how school culture can play a role in teacher perceptions of evaluation systems. Of particular interest in the case of new teacher evaluation policies are the shared assumptions regarding feedback and sharing, since these new systems may introduce new ways of providing feedback and talking about the struggles teachers are facing.

A new teacher evaluation program, especially one that emphasizes accountability, is seen as a potential risk for organizational members. This is because a new evaluation program is a type of organizational change (Fullan, 2007; McLaughlin & Pfeifer, 1988), and organizational members almost always resist change because it poses a threat to the stability of the individual as well as the organization (Evans, 1996). New evaluation policies have implications for compensation, promotion, and dismissal so it is only natural that teachers would respond to these policies with resistance and that these policies would be perceived as threatening. To overcome this perception, a school must have a culture that supports risk-taking, encourages communication, and is built upon trust. In their case studies of districts implementing new teacher evaluation programs, McLaughlin and Pfeifer (1988) found that each successful implementation occurred within a school and district culture that emphasized evaluation as an opportunity for feedback and improvement, rather than a perfunctory exercise to be endured so one can continue to be employed. In these cases, teacher trusted that they would be evaluated fairly, and this influenced the positive perception of the evaluations.

2. **<u>Relationship with Evaluator</u>**

The importance of trust in organizations points to the centrality of the evaluator in shaping teacher perceptions of a new evaluation program. McLaughlin (1990) states, "an effective evaluation system insists on trust between teachers and administrators" (p. 404). Teachers must trust that the evaluators, who are usually administrators, can and will evaluate the teachers fairly. In their set of district case studies, Stiggins and Duke (1988) found that evaluators lacked skills in evaluating teacher performance as well as communicating with teachers regarding the evaluation process (p. 21). If teachers perceive their evaluator to be incompetent or untrustworthy, it is doubtful that they will perceive the evaluation process as legitimate, valid, or reliable. Milanowski and Heneman (2001) also found the competence and skill of the evaluator to be a significant predictor of teacher reactions.

Kelley and Finnegan (2003), in their study of the factors that influence teacher expectancy, found similar results. Teacher expectancy is "the belief that individual effort will result in the achievement of specified goals" (Kelley & Finnegan, p. 604). The teachers in two districts that had recently implemented school wide merit pay systems were asked about the factors that influence their expectancy, or belief that they would be able to earn a bonus. Kelley and Finnegan found that perceived fairness in the system was the single strongest predictor of teacher expectancy. Principal support, feedback, and goal conflict were also significant predictors. Milanowski and Heneman (2001) also found that "fairness concerns also pervaded teachers' perceptions of the new evaluation systems" (p. 209). Thus, the perceived fairness of a teacher evaluation system could also play a role in developing a generally positive or negative perception of the teacher evaluation policy or system.

Often, the evaluator is the principal, which poses complications because this can force the principal to embody two conflicting roles, especially in evaluation systems that blend elements of accountability and professional development (Peterson, 1990). On one hand, the principal must play the role of supporter, encouraging teachers to take risks and continue to improve their practice. This role of growth facilitator has been show to be highly correlated to positive perceptions of evaluations (Duke & Stiggins, 1990). However, principals are also required to play the role of judge, especially as they assess which teachers merit increased compensation, tenure, or even dismissal (Peterson, 1990). Many new teacher evaluation systems partially address this through by mandating that student growth scores play a significant role in evaluations, but observations are also used. Teachers are aware of this dual role, and the way the principal manages the conflicting responsibilities will influence teacher perception of the evaluation system.

3. **Opportunities for Sensemaking**

Finally, teachers are more likely to have a positive perception of new evaluation policies if their organizations give them time and a structured opportunity to consider the new policy and its impacts within the context of their organization (Fullan, 2007; Spillane, Reiser, & Reimer, 2002). Even with the most positive change, organizational members still need dedicated time at their school site to learn and understand the implications of that change. In the case of an accountability-driven teacher evaluation system, that initial learning will be compounded by the need to make sense of a potential threatening or distressing change. Allowing time for this process of meaning making to occur will positively influence teacher perceptions of the programs. As Fullan says, "meaning fuels motivation" (p. 39). If teachers are not given dedicated

time and opportunities to develop their own meanings about a new evaluation system, it is likely that they will lack interest in the new evaluation programs or at worst, continue to feel threatened by them.

D.   **Individual Teacher Characteristics**

This focus on meaning draws attention to a key fact about these programs: they have significant implications for the lives of teachers. Given this, certain factors about individual teachers will shape how they perceive new teacher evaluation policies. Spillane, Reiser, and Reimer (2001) identify individual teacher attributes as the third level of factors that influence sensemaking and perception. They note, "different agents will construct different understandings, seeing what is new in terms of what is already known and believed. What we see is influenced by what we expect to see" (p. 395). Given this, policymakers should not overlook the fact that it is individual teacher and administrators who will implement and be affected by these policies (Evans, 1996; Weatherley & Lipsky, 1977). While a range of individual characteristics could influence how teachers respond to teacher evaluation policies, age and career stage, career goals and professional orientations, group identity, and the relationship between a teacher's identity and the particular policy image embedded in the policy are particularly salient characteristics.

1.   **Career Stage and Age**

Career stage and age influence teacher perceptions of evaluation polices because career stage and age influence teacher motivations. First, new teachers face many of the same problems with evaluation with more veteran teachers, all of which have been discussed above. These

problems include generally poor evaluation practices and the principal as both judge and supporter (Peterson, 1990). However, new teachers are also facing unique demands given the steep learning curve of teaching. Depending upon the culture of the school, new teachers could be disinclined to reach out for assistance when, and it certainly will be, needed. New evaluation policies could be perceived as threatening to these new teachers, as they are now required to perform at high levels to earn tenure, whereas prior to these systems most new teachers were virtually assured of earning tenure.

Evans (1996) exhaustively details the particular circumstances and challenges experiences by midcareer teachers. First, they could be facing burnout, which is disillusionment and a lack of efficacy that leads to loss of caring (Evans, p. 95). Often, veteran teachers have had negative experiences with evaluation in the past (Peterson, 1990). Midcareer teachers may also face a general leveling off of performance and growth in their work. Many midcareer teachers are also in midlife, which often brings new stresses through changing family roles and perhaps a growing focus on material rewards. Certainly, most midcareer teachers continue to derive satisfaction from their work and from student success, it is simply that many teachers have an increased focus on concerns such as salary and benefits (Evans, 1996). Finally, midcareer teachers face the reality that they might receive diminished recognition from their colleagues, as their skills become somewhat taken for granted (Evans, 1996).

2. **<u>Professional Identity</u>**

In addition to age and career stage, a teacher's identity can play a large role in perception of a new teacher evaluation policy. Researchers have defined teacher identity is many ways. Spillane (2000) defines identity as, "an individual's way of understanding and being in the

world" (p. 308), which includes "dispositions, interests, sense of efficacy, locus of control, and orientations towards work and change" (p. 308). Jansen (2001), in his examination of education policy implementation dilemmas, describes teacher identity as "the way teachers feel about themselves professionally, emotionally, and politically given the conditions of their work (p. 242). Sachs (2001) argues, "identity cannot be seen as a fixed 'thing', it is negotiated, open, shifting, ambiguous, the result of culturally available meanings" (p. 154). Additionally, "[identity] provides a shared set of attributes, values, and so on that enable the differentiation of one group from another" (Sachs, p. 154).

Social identity theory argues that teachers derive a portion of their identity from the groups with which they identify (Stets & Burke, 2000; Ashforth & Mael, 1989). Indeed, every individual considers themselves a part of many groups, from a religion, to a political party, or a kickball team. Often, these groups are in conflict and the individual must negotiate how to balance the conflicting beliefs or messages of those groups as part of the process of forming their identity. For example, an individual that is part of the Catholic Church and the Democratic Party must reconcile the conflicting messages those groups have with respect to abortion. In the context of teaching, teachers unions, school communities, and education groups affiliated with corporate-style education reform such as the New Teacher Project, Students First, and Teach for America (TFA) are some of the most prominent examples of groups with which teachers may identify.

In the case of the most recent iteration of teacher evaluation policies, social identity may prove particularly salient. Two of the possible groups with which teachers could identify, teachers unions and corporate-style reform groups, have widely divergent views on accountability-oriented teacher evaluation programs, the use of student achievement data in

evaluations, and the use of evaluations to inform personnel decisions. For example, Students First, an organization founded by TFA alumnae Michelle Rhee, advocates for merit-pay programs and the elimination of tenure (Students First, 2012). In contrast, the National Education Association (NEA), the country's largest teachers union, continues to support tenure, albeit based on performance instead of length of employment (NEA, 2012). Depending on which group, if any, is part of a teacher's identity, that teacher may very well have different perceptions of a new evaluation policy.

Finally, the congruence between a teacher's self-image and the policy image of the teacher embedded in the new teacher evaluation policy could influence perception. Policy image, as defined by Cochran-Smith and Lytle (2006) is the "central common conceptions that are symbolic of basic attitudes and orientations to teaching and learning" (p. 670). Essentially, this is the idea of teachers and teaching that is conveyed through a policy, and the idea of teachers that the policy is built upon. Teachers may perceive policies in different ways, depending on the policy image embedded within the policy. However, teachers may also respond differently based on the fit between their identity and the policy image.

Jansen (2001), in his study of post-apartheid education policy in South Africa, hypothesized that policy images may conflict with personal and professional identities of teachers. This fit, or lack thereof, between self-image and policy image may be a factor that could influence teacher perception. Spillane, Reiser, and Reimer (2002) note that this conflict between policy image and self-image could lead to several different responses. They state:

> Teachers might become advocates by deciding they were "ahead of the curve" and already teaching in ways consistent with the reform. Or they might be motivated to discount the reform idea, seeing it as inconsistent with "the reality" that they "know best." Alternatively, teachers might accept the need for change but attribute the reasons for their not adopting the reform to factors in their context. (p. 403).

Sabatier and Manzmanian (1980) also point to the role this disjunction between self-image and policy image might play in implementation, noting that the more varied a constituencies attitude toward a policy is, the more challenging implementation will be.

In conclusion, a cognitive implementation framework points to the factors that influence the sensemaking process and the formation of teacher perception of new teacher evaluation policies. These factors can be grouped into three categories: policy-level factors, organizational-level factors, and individual teacher level factors. Certainly, these levels all can and do interact, as the policy content and policy signals are filtered through the organizational context to the individual teacher. However, analyzing each level separately allows for a more thorough understanding of how the elements of the policy, the organizational context, and the beliefs, attitudes, and characteristics of the individual teacher may vary from school to school, and teacher to teacher. Combined, these levels provide a complete picture of how teachers can form perceptions of new policies.

III. METHODS

A.   **Research Design**

The development of teacher perceptions of new evaluation policies can be thought of as a process of teachers making sense of a new policy, with its new demands and ramifications for their work. Spillane, Reiser, and Reimer's (2002) cognitive implementation framework provides a set of possible factors that may influence teacher sensemaking, and ultimately perceptions. These factors are grouped into three categories: policy signals, organizational context, and individual teacher characteristics. The first research question, what are teacher perceptions of new evaluation policies and the second research question of this study, what are the factors that influence teacher perception of accountability-oriented evaluation policies, require an investigation into the relationships between possible factors and teacher perceptions of evaluation policy. This investigation into statistically significant relationships and interactions indicates the need for a quantitative approach. The third research question, how are these factors related to overall perceptions of teacher evaluation policies, calls for an in-depth exploration of teachers' sensemaking and engagement with these policies. The need to understand relationships in depth points towards a qualitative component. Both of these components, quantitative and qualitative, make unique contributions to understanding how teachers perceive new evaluation policies.

Accordingly, this study used a mixed methods design (Tashakkori & Teddlie, 2003), which is a research design that incorporates both quantitative and qualitative data, integrating these data types at various points throughout the study. This study design is based on the assumption that "the uses of both quantitative and qualitative methods, in combination, provide a better understanding of the research problem and question than either method by itself"

(Creswell, 2011, p. 535). Quantitative research is often considered explanatory or confirmatory, while qualitative approaches are often considered exploratory. Combining these approaches allows the researcher to perform both types of research within a single study. A mixed methods approach also allows the researcher to provide both "numbers" and "stories", possibly broadening the readership or impact of their research (Creswell, 2011).

Furthermore, a mixed-methods design provides three unique opportunities, when compared to research that is either solely quantitative or qualitative. First, a mixed methods approach can provide better or stronger inferences from the data, because the different data types combine to provide more breadth and depth that either data type would provide on its own (Teddlie & Tashakkori, 2003), especially if care is taken to select methods that have "complementary strengths and non-overlapping weaknesses" (Johnson & Turner, 2003). Second, mixed methods designs allow the researcher to address questions that could otherwise not be addressed (Teddlie & Tashakkori, 2003). For example, combining surveys with interviews allows the researcher to develop a statistical model that explains relationships between variables, while the interviews allow the researcher to explore those relationships in depth. Finally, a mixed methods approach allows for a greater diversity of views, especially divergent views that may point to flaws or incompleteness in existing theory (Teddlie & Tashakkori).

One concern that has been raised regarding mixed methods designs is the possible need to reconcile conflicting paradigms of research. Though the differences between various paradigms have been classified as "wars" (Gage, 1989), further research has shown that the different paradigms of knowledge can be reconciled, or at the least coexist, in a mixed methods study (Greene & Caracelli, 2003). Quantitative research is often associated with positivist or post-positivist paradigms of knowledge, while qualitative research is associated with a constructivist

or relativist approach to knowledge. In their examination of how researchers resolve paradigmatic issues when using mixed methods design, Greene & Caracelli (2003) found that some researchers emphasized the importance of paradigmatic distinctions and embraced a dialectic view, acknowledged that multiple paradigms could make unique contributions to a study, or sought to develop new paradigms. Other researchers minimized the role of paradigms and emphasized a pragmatic approach or made research decisions based on "their ability to further substantive agendas of the inquiries" (p. 103). Regardless of the approach, many avenues exist for researchers to use mixed methods designs without fear of violating research paradigms.

The theoretical perspective is just one assumption implied through different mixed methods designs. Creswell, Clark, Gutmann, and Hanson (2003) identify three assumptions in addition to the theoretical perspective: implementation, priority, and integration. Implementation refers to how the different portions of the study will be conducted. The primary distinction is between concurrent studies, in which the qualitative and quantitative portions occur at the same time, and sequential studies. Priority refers to the relative weight given to the qualitative and quantitative portions. Often, limitations such as time constraints or audience preference force the research to emphasize one type of data over another. Finally, integration refers to the point at which the different data sets will be combined (Creswell, et al., 2003). Combining the data can be done at the end of the study, or at multiple points during the research. By considering these three assumptions – implementation, priority, and integration – the specific design of the study becomes clear.

This study took a concept driven stance, rather than a theoretical perspective that emphasizes the role of research paradigms. A concept driven stance allowed for research decisions to be made in order to advance the understanding of the concept being studied, rather

than their "congruence with particular sets of philosophical assumptions" (Greene & Caracelli, 2003, p. 103). This does not mean that the researcher does not possess philosophical leanings or epistemological commitments that influence how she views the research problem, but rather that the researcher was free to make decisions about design that will deepen the understanding of the concept at hand, even if those decisions cut across different paradigms.

In light of this stance, as well as the research questions, this study used a sequential explanatory design. Spillane, Reiser, and Reimer's (2002) cognitive implementation framework points to three main categories of variables. These categories are the policy signals, the organizational context, and the individual characteristics of the teacher. Within each categories, the possible variables were selected according to the factors that the existing body of literature, which includes sensemaking literature and previous studies of teacher evaluation policies, indicates are likely to influence teacher perception of a new evaluation policy. These variables, outlined below, formed the basis of the quantitative study.

As indicated by the name, the stages of this research were implemented sequentially, with quantitative data collection and analysis completed first, followed by qualitative data collection and analysis. Typically, priority is given to the quantitative results. That was the case in this study, since the major focus was on identifying the factors that could influence teacher perception, as described in the literature. The qualitative results were used to illuminate and expand the findings from the quantitative analysis (Creswell et al., 2003; Creswell, 2011). The two phases were integrated at two points. The results from the quantitative piece were used to provide context for the interviews, and the analysis of each phase was integrated during the interpretation of the results. One of the main strengths of this design was its clear-cut stages, as well as the way it lent itself to distinct analysis and reporting sections. The primary drawback

was the length of time it takes, since data collection and analysis of the quantitative piece must often be completed prior to beginning data collection on the qualitative portion of the study (Creswell, et al., 2003).

The nature of the sequential explanatory design also aligned with the research questions of this study. The first question, what are teacher perceptions of evaluation, and the second research question, what factors influence teacher perception of new accountability-based teacher evaluation policies, were addressed through the quantitative part of the study. The third question, how are these factors related to overall perception of teacher evaluation systems, was addressed in the qualitative portion of the study. Together, these answers to these two questions provided both a broad overview of the relationships between different factors and perception, as well as a deeper look into the sensemaking and perception development of teachers when faced with new evaluation policies.

B.   **Variables in the Quantitative Analysis**

The variables in the quantitative analysis were drawn from the existing literature on teacher perceptions of evaluation policy. Following Spillane, Reiser, and Reimer's (2002) cognitive implementation framework, the predictor variables are grouped into three categories: policy signals, organizational context, and individual teacher characteristics, all of which may influence the sensemaking process and development of teacher perceptions. The outcome of interest, teacher perception of new evaluation policies, was measured using three constructs: satisfaction, purpose, and impact. Each concept was measured using five items on the survey. These items were used in exploratory factor analysis to develop perception constructs, which increased the reliability of the constructs and increased the likelihood of a normal distribution for

the outcome variables. In turn, a normal distribution of perception constructs allowed for parametric analysis, rather than a non-parametric approach.

For the construct 'satisfaction', teachers were first asked, "In general, are you satisfied with TES?" and responded by selecting a response ranging from Strongly Disagree to Strongly Agree. To measure other aspects of satisfaction, teachers were asked to respond to four other statements asking about their satisfaction with the new teacher evaluation system. The response for each of these items was measured using a 5-point Likert scale. Likewise, the construct 'impact', was measured using an item that asks teachers to describe the effect the new evaluation system had on their practice, followed by a series of four statements about the impact of the system on professional growth, feedback, quality of instructional conversation, and student performance. Again, participants indicated their response using a 5-point Likert scale. Finally, the construct 'purpose' was measured using 5-point Likert scale questions asking teachers to rate the perceived purposes of the evaluation policy.

The predictor variables were drawn from a review of the relevant literature on teacher evaluation policy and teacher perceptions. The literature suggests that these factors may influence how teachers perceive new policies, and are therefore appropriate factors to include when addressing the first and second research questions. While a large number of variables may influence teacher perception, the variables that the literature indicates are most likely to influence perception were selected. Grouped into policy, organization, and individual levels, the factors are as follows:

Policy Level Factors

- Elements of the evaluation system: standards used in the evaluation system, perceived validity of the instruments in the evaluation system

- Communication: where teachers received information about the evaluation system

- Images of teaching and teachers: conception of teacher embedded in the system, policy images of teachers implicit in the evaluation system

- Purpose of the evaluation system: what do teachers think this system is designed to do

Organizational Level Factors

- Culture: the extent to which the organizational culture is open to feedback and accountability

- Sensemaking: the extent to which there are opportunities for teachers to develop shared meanings around the new evaluation system,

- Evaluator: trust in the evaluator, credibility of the evaluator, perceived fairness

Individual level factors

- Career stage and age

- Identity: social identity and alignment between identity and policy image

Demographic characteristics such as race and ethnicity and gender were used as control variables. Table 1 shows the relationships between the constructs that were measured and the survey items. Additionally, the appendix contains the full survey, with each item marked to show the construct it measured.

Table I

Factors and Predictor Variables

Level	Factor	Construct	Survey Item
	Perception	"Satisfaction"	3-7
		"Impact"	8-12
Policy Level	Related to the purpose of the evaluation system	"Purpose"	13-17
	Related to elements of the evaluation system	"Perceived Validity"	19-22
	Related to communication of the policy	"Information source"	18
		"Policy Image"	23-29
	Related to images of teaching and teachers	"Culture"	43-47
Organizational Level	Related to the organizational context	"Sensemaking"	30-34
		"Evaluator"	36-42
Individual Level	Related to identity	"Disposition"	48-50
		"Affiliation"	59
		"Certification"	60

C. **Quantitative Sampling**

The target population in this study was all fulltime K-12 teachers in River District. River

District is a pseudonym for a large suburban district in Illinois. This district provided an

opportunity to examine how teachers view changes to teacher evaluation policies in response to

federal pressure and state legislation in a large suburban district. Due to the unique

characteristics of the teacher evaluation system in River District, the population of this study is

limited to this district, and cannot be extended to teachers in other districts in Illinois. Teachers

who met the following criteria could participate in the study: (1) full-time teacher in River

District during the 2013-2014 school year. Part-time teachers will be excluded because they do

not have the same work requirements, in terms of hours or number of students, as traditional full-

time teachers.

The quantitative phase of the study used a convenience sample, which consisted of approximately 85% of the River District teachers who shared their email address with the teachers union. A convenience sample makes it unlikely that any conclusions or inferences drawn from the study are generalizable to the population, but it provided useful information that addressed the research questions (Creswell, 2012).

D.   **Qualitative Sampling**

The qualitative portion of the study used a convenience sample, derived from the survey participants who indicated they were interested in being interviewed and responded to interview requests. Spillane, Reiser, and Reimer's (2002) cognitive implementation framework indicates that policy signals, organizational context, and individual characteristics are all factors which are likely to influence how teachers make sense of and view new evaluation policies. Furthermore, they argue that these factors may interact within and across levels as teachers construct their perceptions. The quantitative analysis showed which factors were significant predictors of how teachers perceived the impact of new evaluation policies and their satisfaction with these policies, but it was not able to present a more in-depth picture of how these factors are related to overall teacher perceptions of the evaluation system.

In order to better understand how factors identified through the survey were related to the overall perception of the evaluation system, follow-up interviews were conducted with 5% of the respondents to the survey. A review of the literature supported the following broad expectations:

- Teachers who think the evaluation system will be used to remove teachers will have a more negative overall view of the evaluation system (Stiggins & Duke, 1988)

- Teachers who do not see the evaluation measures as valid will have a more negative view of the evaluation system (Bacharach, Conley, & Shedd,1990; Peterson & Comeaux, 1990).

- Teachers who do not think the evaluation system reflects the professional nature of their work will have a more negative overall view of the evaluation system (Peterson & Comeaux, 1990; Wise, Darling-Hammond, McLaughlin, & Bernstein, 1985).

- Teachers whose school culture is more open to sharing and feedback will have a more positive view of the evaluation system (Firestone & Louis, 1999; Schein, 1996).

- A more positive relationship with the evaluator will be related to a more positive view of the evaluation system (Duke & Stiggins, 1990; Milanowski & Heneman, 2001).

- Career stage and career length will be related to a more negative view of the evaluation system (Evans, 1996).

The goal of this study was to provide a cohesive picture of teacher perceptions of new evaluation policies, as well as the factors that are significantly related to those perceptions. The existing research on teacher evaluation, as well as the cognitive implementation framework, point towards likely factors. By identifying the factors and perceptions, the survey portion of the study addressed the first two research questions: what are teacher perceptions of new evaluation systems and what factors are significantly related to those perceptions? The interview portion of the study addressed the third research question, what are the relationships between the identified

factors and overall teacher perceptions, by asking selected teachers to expand on their

impressions of the evaluation system.

E.   **Phase I Quantitative**

    1.   **Data Collection**

The first phase of the study focused on identifying the relationships between the variables

listed above and the outcome variables related to perception. Data were collected using a cross-

sectional survey, with data being collected at only one point in time (Creswell, 2011). The

questionnaire was a combination of self-developed items, items from the Consortium on Chicago

School Research's My School, My Voice Survey (CCSR, 2012), and Stiggins and Duke's (1988)

Teacher Evaluation Profile (TEP) questionnaire. When possible, existing items were used but

selected variables, such as teacher identity, required the development of new items. A variety of

item types were used, including Agree/Disagree questions measured on a 5-point Likert scale,

multiple-choice questions, and ranking questions. The survey was organized into four sections.

The first section contained the items designed to measure the policy level variables, the second

addressed the organizational variables, and the third measured the individual level variables. The

fourth section collected demographic data, which was used as control variables.

The survey was distributed using via email as a Web-based survey and used Qualtrics as

the survey software. All emails came from the president of the teachers union, in order to protect

the email addresses and identities of the respondents by not sharing contact information with the

researcher. First, an email introducing the study, explaining its purpose, and containing the link

to the Web-based survey was sent out to the email addresses provided to the president of the

teachers union.-Though previous studies have found that Web-based surveys have a lower

response rate and coverage than mail surveys (Jacob, 2011; Converse, Wolfe, Huang, & Oswald, 2008), the logistical issues of using paper surveys for over 2000 teachers with limited research resources made Web-based surveys the best choice for this study. After one week, a reminder email was sent to all respondents, even if they had completed the survey. A second email reminder was sent one week after that, reminding participants of the importance of their input.

To encourage participants to complete the survey, a raffle was used to award three $50 Amazon gift cards to three randomly selected survey participants. After the second and last reminder email was sent out, an email letting participants know that if they complete the survey they were eligible for a raffle was also sent out. Participants who had already completed the survey were directed towards a survey where they can enter their name and email address for the raffle. A link to this survey was also added at the end of the study survey, and participants who completed the study survey after receiving the raffle email could enter their information upon completing the survey. The information entered for the raffle drawing was not linked to the survey responses in any way. Participants did not have to enter the raffle to participate in the study; it was simply an incentive to encourage participation. Raffle winners were selected using a random number generator two weeks after the raffle email was sent out.

2. **Data Analysis**

The second research question, what are the factors that influence teacher perception of evaluation policies, informed the choice of analysis. Multiple regression was an appropriate choice to address this question, because it showed the direction and magnitude of the relationships between predictor variables and teacher perceptions of the evaluation policy. One

multiple regression models was used, which used the overall perception construct as the outcome variable.

Prior to any multivariate or regression analysis, descriptive statistics for all of the variables were reported. Frequency tables for each item were also reported, in order to examine the frequencies of each item response. The data was also screened for accuracy, to ensure that the means and standard deviations for each variable were plausible. Missing data was also examined to see if the data was missing completely at random, which implied the respondent forgot to answer a particular question, at random, or not at random. A t-test between data with missing values and data with non-missing values was conducted to identify the type of missing data. Upon examining the missing data, it was clear that the missing responses were due to participants failing to complete the survey. SPSS was able to handle this type of missing data, so no modifications were made to the data set. Data was also screened for outliers using z-scores and histograms.

Once descriptive statistics for all items were examined and the dataset was screened for missing data, a factor analysis was used to reduce the number of predictor variables prior to a multiple regression analysis. Factor analysis is a process of estimating the underlying factors, or latent constructs, in a set of variables, while also removing the possibility of multicollinearity between the variables. Once the factor loadings were determined, it was possible to calculate individual scores for each construct. These individual teacher scores on each construct formed the dataset used for the multiple regression analysis. This allowed the model to have fewer predictors and predictors that were more likely to meet the underlying assumptions for regression analysis.

Once the factor loadings were determined and the individual scores for each construct were calculated, the data were tested to see if they met the assumptions of multiple regression. First, the data were tested for normality, using measures of skewness and kurtosis, as well as a histogram. Data were tested for linearity, using scatter plots and residual plots. The assumption of homoscedasticity will also be tested, using scatter plots. Data were tested for multicollinearity and singularity. Multicollinearity was addressed during the factor analysis, but the factor loadings were combined with control variables such as age and race in the regression analysis, so the control variables must also meet the assumptions.

3.   **Reliability and Validity**

To confirm the reliability of the survey instrument, cognitive pretesting was used. Cognitive pretesting consists of asking the respondent to "think aloud" while the questionnaire is completed (Krosnick, 1999), in order for the researcher to identify any issues with phrasing and to confirm that the items are understood as intended. Cognitive pre-tests were conducted with two teachers and one district official at River District. The three pre-testers read through the survey with the researcher, and identified issues with terminology or points of confusion as they arose.

Additionally, many of the items on the questionnaire have been rigorously tested as part of other research projects. Many survey items are drawn from the Consortium for Chicago School Research's My Voice, My School Teacher Survey, which is an annual survey given to teachers, parents, and students in Chicago Public Schools. Stiggins and Duke (1988) developed their Teacher Evaluation Profile questionnaire after a series of case studies about teacher evaluation in a wide variety of districts. Finally, Cronbach's Alpha was reported for each set of

items designed to measure a particular construct. This provided an indicator of internal consistency before moving forward with the analysis.

Finally, the use of factor analysis also served as validation of the survey instrument. As shown in Table I, the survey was designed so that several items measured each hypothesized factor. A factor analysis of these items showed that the majority of the items, especially for the constructs 'Evaluator', 'Culture', and 'Sensemaking', loaded onto their predicted factor. This confirmed that the items on the survey reflected the construct they were designed to measure.

F.   **Phase II Qualitative**

1.   **Data Collection**

The purpose of the qualitative phase of the study was better explain the results from the quantitative phase, and to address the third research question, how are the identified factors related to overall perceptions of new teacher evaluation policies. The interviews were semi-structured and conducted in person at a site of the participant's choosing, often their school site or a coffee shop. Interview participants consisted of survey respondents who indicated a willingness to be interviewed and responded to an interview request. The last page of the survey had a section where teachers indicated if they are interesting in being contacted for a follow-up interview.

A copy of the full interview protocol can be found in the Appendix. The interview began with the question, "Tell me, in general, what do you think of TES?". Next, the participants was asked " What do you think TES is designed to accomplish?", as well as how it will impact their desire to continue teaching in their district, their desire to remain a teacher, and how they think the evaluation system affects student learning. In the next section, teachers were asked about

their prior experiences with evaluation and how they think their relationship with their evaluator will affect their evaluation process. They will also be asked about how district officials, the teachers union and other individuals may have influenced their perception of the teacher evaluation system. Participants were also asked, "Why do you think there is a policy emphasis on teacher evaluation right now" and "Why do you think most new teacher evaluation systems incorporate test scores". They were also asked about the sources they used to learn about the teacher evaluation system in their district.

Finally, teachers were asked how they would design their own evaluation system, and if they think teachers need to be evaluated at all. These questions gave teachers an opportunity to talk about the parts of teaching that they think are most important. Also, by talking about how they think teachers should be evaluated, this created an opening to talk about what they believe the work of teaching and may reveal misalignment between the parts of teaching the teachers value, and the parts of teaching that the district teacher evaluation system values.

All data collected through the interview portion of the study will be stored in encrypted files on the investigator's laptop for three years after the completion of the study. All interviews were audio-recorded with participant consent. Those audio recordings are password-protected and stored on the investigator's computer. All transcribed interviews use pseudonyms for the participants and their schools, and any other identifying information was removed.

2.   **Data Analysis**

In the second phase, data analysis was an iterative process, which allowed for the incorporation of new themes and ideas over the course of the analysis. Overall, the coding process followed five steps (Creswell, 2011). First, all transcribed interviews were read as part of

a preliminary exploratory analysis in order to develop a general sense of the data. Next, the transcripts were broken into smaller segments of information. An initial set of codes was developed based upon the theoretical framework and review of the literature. A code was developed for each potential factor listed in Table I, as well as the new factors generated by the factor analysis. The coding was an iterative process and new codes were developed as needed. If the researcher identified an idea or theme, such as the role of rumor, in two or more transcripts while coding, a new code was created and all transcripts were read through again with this code in mind. In this way, new codes were added to the initial, literature-based, set of codes.

Next, once all interviews were coded, the final set of codes was checked for redundancy to see if the total number of codes could be reduced. Thirty-six unique codes were identified. The codes were used to generate themes and descriptions of the identified factors were related to overall teacher perceptions. To do this, the number of times each code was used to identify the most common themes mentioned in the interviews. Since each piece of text could be assigned multiple codes, the co-occurrence of codes was also examined to look for codes that were commonly applied together. These two metrics provided the researcher with a sense of the most common codes, as well as codes that frequently appeared together. Next, the coded pieces of texts were compared across interviews for each code and groups of code. The researcher noted commonalities and differences across the interviews. As this was done for each code, three different groups based on overall perceptions began to emerge. Once these three groups were apparent, the coded excerpts for each code were also compared across groups. This allowed the researcher to examine how each group's responses varied according to each theme.

A.   <u>Summary</u>

Based on the findings from this study, a professional approach to teacher evaluation in new evaluation systems is crucial to positive teacher perception of those systems. The belief that the evaluation system supported teacher professionalism, along with the teachers' beliefs about their evaluator and their perceived familiarity with the evaluation, were significant predictors of overall perception. The interview portion of the study revealed that teachers were more likely to have positive opinions of the teacher evaluation system when they believed core components of the system, including the observational rubric, aligned with their beliefs about teaching and conveyed a professional image of teaching. Additionally, teachers who felt their evaluators viewed them as competent professionals and operated from a shared understanding of the goals of the evaluation system were more likely to have a positive opinion of the evaluation system.

B.   <u>Quantitative Findings</u>

1.   <u>Introduction</u>

The survey portion of the study provided an overall picture of how teachers in River District view the teacher evaluation system, as well as the factors that are significantly related to teacher perceptions of the evaluation system. Overall, teachers in River District had a neutral perception of the teacher evaluation system [TES], neither seeing it as harmful nor helpful. Several factors, including teacher professionalism, knowledge of the evaluation system, and the evaluator, were significantly related to teacher perceptions of the evaluation system. The results from the survey portion of the study highlight the relationship between teacher belief in the ability of a new teacher evaluation system to boost the professionalism of the teacher workforce

and overall positive perceptions of the evaluation system. Additionally, it draws attention to the need for increased communication and training on the requirements and uses of the evaluation system, since teacher knowledge of the evaluation system is positively related to overall perceptions of the evaluation system. Furthermore, a teacher's relationship with his or her evaluator is often cited as a crucial factor in teacher perception of an evaluation system, which proved to be the case in this study.

2.   **<u>Sample Characteristics</u>**

In this study, the response rate was 9.46%, with 219 respondents. Of the respondents who chose to provide their demographic data, 20.1% of the respondents identified as male and 79.9% identified as female. When asked about their tenure statues, 90.1% of the respondents indicated they have tenure, while 9.9% of the respondents indicated they do not. The reported age ranged from 29 to 68, with a mean age of 47. The reported race and ethnicities of participants are shown in Table II. Both the racial and ethic breakdown, as well as the proportions of men and women are similar to characteristics of the teaching force as a whole.

Table II

Reported Race/Ethnicity of Respondents

Race/Ethnicity	n	Percent
White	162	85.3%
Spanish/Hispanic/Latino	21	11.1%
Black or African-American	4	2.1%
American Indian or Alaskan Native	0	0.0%
Asian	1	0.5%
Middle Eastern	1	0.5%
Multi Ethnic	1	0.5%

Note. Categories are taken from U.S. Census.

River District serves pre-kindergarten students through high school. The grade levels of the

sample are shown in Table III. It should be noted that in this district, sixth grade is taught within

elementary schools and middle schools only serve seventh and eighth graders.

Table III

Reported Grade Level Taught by Respondents

Grade Level	n	Percent
Elementary	86	47.0%
Middle	44	24.0%
High School	53	29.0%

Respondents were also asked to describe their political orientations, since teacher evaluation

policy is often politicized or presented differently by those with varying political affiliations.

Table IV shows the general political orientations of the respondents.

Table IV

Respondent Political Orientations

Orientation	n	Percent
Extremely Liberal	6	3.2%
Liberal	42	22.2%
Slightly Liberal	20	10.6%
Moderate/Middle of the Road	61	32.3%
Slightly Conservative	33	17.5%
Conservative	25	13.2%
Extremely Conservative	2	1.1%

Note. Categories are taken from the ANES Public Affairs Survey.

Finally, the majority of respondents have earned a masters degree. Table V shows the highest

degree earned.

Table V

Reported Highest Degree Earned

Degree	n	Percent
Bachelors Degree	22	11.0%
Masters Degree	165	82.5%
Professional Degree (MD, DDS, DVM, JD, DD)	8	4.0%
Doctorate (PhD or EdD)	5	2.5%

3.   **<u>Factor Analysis</u>**

Exploratory factor analysis was used to reduce the number of predictor variables prior to regression analysis. Factor analysis is a process of estimating the underlying factors, or latent constructs, in a set of variables, while also removing the possibility of multicollinearity between the variables. Exploratory factor analysis was used, rather than confirmatory factor analysis, because while there is broad literature base around teacher evaluation, there are not strong theoretical or empirical models to form the basis for confirmatory factor analysis. Rather, exploratory factor analysis was used to look for latent structures in the collected data.

Prior to factor analysis, the data were screened for accuracy, by checking that the mean and range for each item was plausible. Since almost all survey items were closed response, no issues were found regarding the accuracy of the data. The data were also screened for missing data. When missing data were found, it was clear that the data were missing because the respondent did not complete the survey. Items were not skipped and there was no discernable pattern to the missing data, other than simple failure to complete the survey. Thus, the data can be thought of as missing at completely at random, and SPSS was able to handle the missing data during analysis without the need to remove all responses from participants who did not complete the survey.

The item of interest in this study was teacher perception of new teacher evaluation policies. In this study, the two key aspects of perceptions are the extent to which teachers are satisfied with the teacher evaluation system in their district, and the extent to which teachers think the evaluation system is changing teaching and learning. To measure this, two constructs were developed: satisfaction and impact. Teachers were asked about how satisfied they were with the teacher evaluation system and if they believe the teacher evaluation system had a positive or negative impact on teaching and learning in River District. The satisfaction construct was measured using five Likert scale items. The impact construct was also measured using five Likert scale items. Table VI summarizes the responses to these items.

Table VI

Survey Responses to Perception Items

Variable	n	M	SD	Range	Skew	Kurtosis
Satisfaction						
In general, I am satisfied with TES	219	3.15	1.02	1-5	0.16	-0.83
TES can be fairly applied to all teachers	219	2.78	1.15	1-5	0.07	-1.10
TES addresses almost all important aspects of teachers work	218	3.04	1.17	1-5	-0.30	-1.06
TES increases my satisfaction with working in my school	219	2.39	0.97	1-5	0.35	-0.48
TES increases my satisfaction with working in my district	219	2.39	1.01	1-5	0.41	-0.38
Impact						
TES supports my professional growth	218	3.12	1.11	1-5	-0.32	-0.95
TES improves the feedback I receive from my principal	219	2.82	1.13	1-5	0.05	-0.93
TES improves the quality of instructional conversations	219	2.87	1.10	1-5	-0.01	-0.91
TES leads to improved student performance	219	2.52	1.10	1-5	0.42	-0.53
TES helps me improve my teaching	218	2.89	1.15	1-5	0.06	-0.92

Next, exploratory factor analysis was done to reduce the number of dependent variables and look for latent structure in the data. As shown in Table VII, all ten items loaded onto one factor, which is overall teacher perception of the teacher evaluation system (α = .94). Rotation was not possible because there was only one factor. Since all the items loaded onto one factor, this factor was used as the dependent variable of interest. Rather than measuring 'impact' or 'satisfaction' separately, these factor is the overall teacher perception of the evaluation system.

Table VII

Factor Loadings for Exploratory Factor Analysis of Teacher Perception Scales

Scale	Overall Teacher Perception
In general, I am satisfied with TES	.81
TES can be fairly applied to all teachers.	.67
TES addresses almost all important aspects of teachers' work.	.68
TES increases my satisfaction with working in my school.	.84
TES increases my satisfaction with working in my district.	.84
TES supports my professional growth.	.81
TES improves the feedback I receive from my principal.	.76
TES improves the quality of instructional conversations.	.76
TES leads to improved student performance.	.83
TES helps me to improve my teaching.	.87

Note. Rotation was not possible because there was only one factor.

The overall teacher perception component (*M*=0, *SD* =0.97) is reliable regardless of sample size because all ten loadings are above .6 and six are above .8 (Guadagnoli & Velicer, 1988).

Figure 2 shows the frequency distribution of the overall teacher perception factor scores.

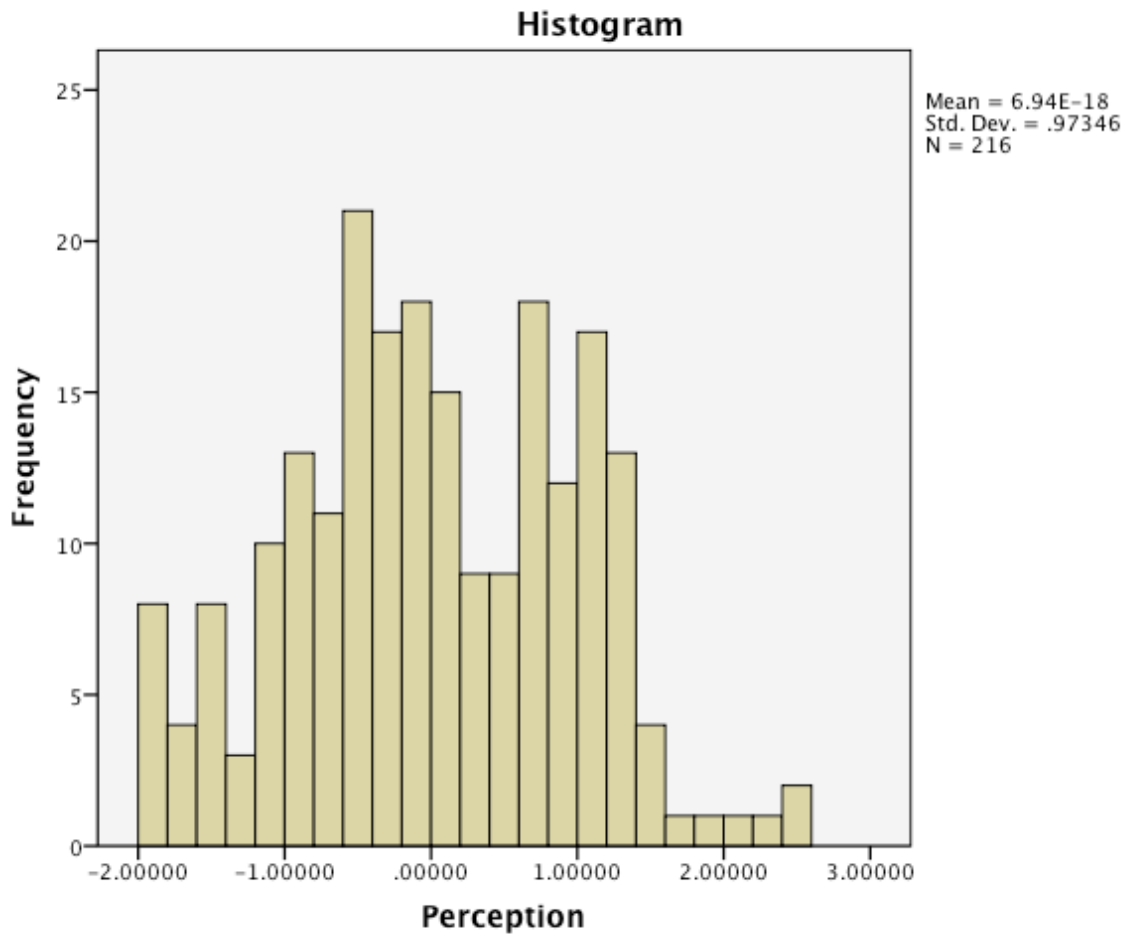### Histogram

Mean = 6.94E−18
Std. Dev. = .97346
N = 216

Figure 2. Frequency distribution of overall teacher perception scores.

Exploratory factor analysis was used to reduce the number of independent variables. Table VIII provides descriptive statistics for all items used in the subsequent factor analysis.

Table VIII

Descriptive Statistics for Independent Variable Survey Items

Variable	n	M	SD	Skew	Kurtosis
TES is used to help struggling teachers.	219	3.04	0.98	-0.37	-0.82
TES is used to raise student achievement.	218	2.53	0.99	0.16	-0.79
TES is used to improve the quality of schools in my district.	219	2.52	0.99	0.11	-0.93
State standardized test scores will accurately reflect student learning.	218	1.81	0.82	0.97	0.68
State standardized test scores will accurately reflect teacher performance.	218	1.61	0.74	1.18	1.20
The TES Framework is a valid and accurate way of assessing teacher quality.	219	2.65	1.15	0.01	-1.07
The TES Framework addresses almost all aspects of teachers' work.	218	2.67	1.23	0.12	-1.22
TES recognizes the professional nature of teaching.	218	3.13	1.12	-0.49	-0.71
TES includes the aspects of teaching that I think are the most important	217	3.10	1.01	-0.45	-0.74
Overall, state standardized tests are good indicators of teacher quality.	218	1.61	0.90	1.87	3.80
Teachers are the most important in-school factor in student success.	218	3.89	1.07	-0.95	0.26
Teachers are the most important overall factor in student success.	217	2.94	1.27	0.12	-1.20
My school leader(s) clearly and carefully explained how the state teacher evaluation requirements will impact TES at my school.	218	3.20	1.68	0.13	-1.53
My school leader(s) gave me and other teachers an opportunity to ask questions about changes to TES.	217	3.59	1.67	-0.35	-1.46
I feel like I will be involved in the implementation of the TES changes at my school.	218	2.74	1.51	0.58	-1.04
I have had all of my questions about the state teacher evaluation requirements answered.	217	3.07	1.51	0.09	-1.32
I had enough opportunities to learn about how the state teacher evaluation requirements will affect me.	218	2.97	1.50	0.28	-1.30
My evaluator is a credible source of feedback.	204	3.28	1.19	-0.62	-0.69
I have a cooperative working relationship with my evaluator.	203	3.64	1.13	-0.90	0.06
I trust my evaluator.	204	3.25	1.31	-0.38	-1.00
My evaluator has a supportive interpersonal manner.	204	3.32	1.24	-0.50	-0.85
My evaluator is knowledgeable about the technical aspects of teaching.	204	3.29	1.18	-0.61	-0.61
My evaluator is knowledgeable about my particular classroom.	204	2.93	1.26	-0.16	-1.19
My evaluator is fair.	203	3.40	1.11	-0.54	-0.32
The culture of my school encourages teachers to share the struggles they have in their classrooms.	204	3.06	1.31	-0.27	-1.21
Teachers talk about instruction in the teachers' lounge, faculty meetings, or other locations in the school.	204	3.63	1.10	-0.97	0.28
Teachers in this school share and discuss student work with other teachers.	204	3.60	1.08	-0.86	0.14
Experienced teacher are willing to share their practice with new teachers.	204	3.93	1.01	-1.21	1.31
Teachers consider evaluation as an opportunity to get feedback on our teaching.	203	2.76	1.11	0.05	-0.95

Exploratory factor analysis with an oblique rotation of the 29 Likert scale items was conducted.

The Kaiser-Meyer Olkin measure of sampling adequacy suggested that the sample was

factorable (KMO = .90). Oblique rotation was appropriate given the correlations between the

majority of the factors, as shown in Table IX.

Table IX

Factor Correlation Matrix for Promax Rotation

Factor	1	2	3	4	5
1	1.00	**0.46**	**0.47**	**0.44**	0.21
2	**0.46**	1.00	**0.43**	0.30	**0.39**
3	**0.47**	**0.43**	1.00	0.31	0.32
4	**0.44**	0.30	0.31	1.00	0.16
5	0.21	**0.39**	0.32	0.16	1.00

Note. Correlations > 0.32 are in boldface.

Five factors emerged from the analysis. The rotated factor loading matrix is presented in Table
X.

Table X

Factor Loadings for Exploratory Factor Analysis with Promax Rotation

| | | | Factor | | |
Item	1	2	3	4	5
My evaluator is a credible source of feedback.	**0.88**	0.38	0.46	0.34	0.21
I have a cooperative working relationship with my evaluator.	**0.86**	0.43	0.39	0.36	0.14
I trust my evaluator.	**0.92**	0.45	0.35	0.39	0.20
My evaluator has a supportive interpersonal manner.	**0.89**	0.40	0.32	0.34	0.18
My evaluator is knowledgeable about the technical aspects of teaching.	**0.75**	0.33	0.42	0.31	0.16
My evaluator is knowledgeable about my particular classroom.	**0.82**	0.34	0.44	0.36	0.20
My evaluator is fair.	**0.89**	0.38	0.39	0.39	0.15
TES is used to help struggling teachers.	0.48	**0.66**	0.28	0.26	0.18
TES is used to raise student achievement.	0.38	**0.75**	0.42	0.23	0.40
TES is used to improve the quality of schools in my district.	0.43	**0.77**	0.37	0.29	0.34
The TES Framework is a valid and accurate way of assessing teacher quality.	0.30	**0.82**	0.28	0.21	0.28
The TES Framework addresses almost all aspects of teachers' work.	0.30	**0.76**	0.29	0.21	0.28
TES recognizes the professional nature of teaching.	0.42	**0.85**	0.39	0.22	0.34
TES includes the aspects of teaching that I think are the most important	0.27	**0.82**	0.31	0.18	0.29
Teachers consider evaluation as an opportunity to get feedback on our teaching.	0.46	**0.56**	0.44	0.48	0.31
My school leader(s) clearly and carefully explained how the state teacher evaluation requirements will impact TAP at my school.	0.41	0.26	**0.80**	0.30	0.18
My school leader(s) gave me and other teachers an opportunity to ask questions about changes to TES.	0.41	0.32	**0.81**	0.27	0.20
I feel like I will be involved in the implementation of the TES changes at my school.	0.44	0.39	**0.65**	0.28	0.30
I have had all of my questions about the state teacher evaluation requirements answered.	0.31	0.36	**0.86**	0.17	0.30
I had enough opportunities to learn about how the state teacher evaluation requirements will affect me.	0.31	0.37	**0.81**	0.16	0.25
The culture of my school encourages teachers to share the struggles they have in their classrooms.	**0.56**	0.36	0.42	**0.62**	0.24
Teachers talk about instruction in the teachers' lounge, faculty meetings, or other locations in the school.	0.36	0.22	0.28	**0.89**	0.16
Teachers in this school share and discuss student work with other teachers.	0.32	0.18	0.22	**0.86**	0.08
Experienced teacher are willing to share their practice with new teachers.	0.22	0.20	0.09	**0.70**	0.07
State standardized test scores will accurately reflect student learning.	0.17	0.27	0.25	0.09	**0.76**
State standardized test scores will accurately reflect teacher performance.	0.16	0.33	0.27	0.10	**0.93**
Overall, state standardized tests are good indicators of teacher quality.	0.07	0.20	0.10	0.13	**0.50**

Note. Factor loadings > .50 are in boldface. Extraction Method: Principal Axis Factoring. TES stands for Teacher Evaluation System.

Seven items loaded onto Factor 1 (α = .95). It was clear that these items were all related to the respondents' beliefs about their evaluator. Accordingly, Factor 1 was labeled "Beliefs about Evaluator". Eight items loaded onto Factor 2 (α = .91). These items all shared a focus on the extent to which the teacher evaluation system was seen as raising teacher quality through an emphasis on teacher professionalism. Factor 2 was labeled "Teacher Professionalism". Five items loaded onto Factor 3 (α = .88). These items all measure the extent to which teachers felt they had ample opportunities for sensemaking and developing new understandings of changes to the teacher evaluation system in response to new state requirements. Factor 3 was labeled "Opportunities for Sensemaking". Four factors loaded onto Factor 4 (α = .84). These items are related to the culture of the school, and the extent to which teachers talk about their practice, share struggles, and mentor newer teachers. Factor 4 was labeled "School Culture". Finally, three items loaded onto Factor 5 (α = .78). These items are related to teacher beliefs about the role of state standardized tests in measuring teacher quality. Factor 5 was labeled "Beliefs about Standardized Tests".

Factor scores were saved for each identified factor, as well as for the dependent variable factor. Descriptive statistics for each composite score are presented in Table XI.

Table XI
Descriptive Statistics for EFA Composite Scores

Factor	*M*	*SD*	Range	Skew	Kurtosis
Evaluator	3.30	1.06	4.00	-0.46	-0.57
Teacher Professionalism	3.38	0.84	3.38	-0.29	-0.78
Culture	3.56	0.92	4.00	-0.87	0.60
Sensemaking	3.12	1.30	5	0.03	-0.83
Beliefs about Standardized Tests	2.02	0.69	3.00	0.41	-0.18

4.   **<u>Regression Analysis</u>**

Using the saved factor scores, control variables, and other data points collected from the survey, multiple regression analysis was used to investigate which variables were significant predictors of teacher perception of the teacher evaluation system, as well as the direction and magnitude of those predictors. Backwards regression was used to order to investigate the full model and then remove variables that were not significant predictors in order to develop the most parsimonious model, and identify the variables most consistently related to the dependent variable (Fowler & Walberg, 1991). The criteria for retention in the model was $p >$ .10, however only predictors with $p > .10$ were considered significant predictors. Interaction terms between beliefs about evaluator, teacher professionalism, opportunities for sensemaking, school culture, and beliefs about standardized testing were included. Control variables were race, age, gender, grade level taught, highest degree, and political orientation. Since the model generated by the backwards regression model indicated that interactions terms were significant but dropped the lower order terms, these terms were added back into the model and the model was run again. In this model, the lower order terms were significant, but the interaction terms were no longer significant.

Regression results indicate that the overall model significantly predicts teacher perception of the teacher evaluation system, $R^2 = .784$, $R^2_{adjusted} = .791$, $F(3,190) = 229.8$, $p = .000$. This model accounts for 79.1% of variance in teacher perception of the evaluation system. A summary of the final model can be seen in Table XII.

Table XII

Predictors of Overall Teacher Perception

Predictor	B	S.E.	β	*p*
Constant	-0.31	1.76		.080
Teacher Professionalism	0.83**	0.04	0.83**	.000
Evaluator	0.09*	0.04	0.07*	.029
Familiarity with TES	0.10*	0.05	0.07*	.057

Note. N=171. *$p <.1$, **$p < .01$

The underlying data meet the assumptions for linear regression. The data were tested by plotting

the unstandardized residual versus the unstandardized predicted value of overall teacher

perception, and verifying that the loess line generally follows the zero-line. Additionally, the

standardized regression residuals are normally distributed.

5. **Factors Influencing Perception of Teacher Evaluation Systems**

The survey portion of the study was designed to address the first two research questions.

The first research question was: what are teacher perceptions of new accountability-oriented

teacher evaluation systems? The survey distributed to all members of the teachers union in River

District revealed that overall teacher perception of the evaluation system is neutral, in which

teachers do not see the evaluation system as a force for improving their instruction, but also do

not see it as harmful to teaching or learning. This neutral view of the teacher evaluation in River

District is somewhat surprising given the rhetoric and tone of the discussion about teacher evaluation policy nationally and in Illinois.

An examination of the survey responses to the overall perception survey items shows that mean teacher responses to some overall perception items were slightly positive, with means just above three, while mean teacher responses to others were slightly negative, with means just below three. The response to the item 'In general, I am satisfied with TES' ($M = 3.15$, $SD = 1.02$) and the response to the item 'TES supports my professional growth' ($M = 3.12$, $SD = 1.11$) were the most positive responses. The response to the item 'TES increases my satisfaction with working in my schools' ($M = 2.39$, $SD = 0.97$) and the response to the item 'TES increases my satisfaction with working in my district' ($M = 2.39$, $SD = 1.01$) were the most negative responses. Additionally, while the survey was designed to measure a 'satisfaction' construct and an 'impact' construct, a factor analysis showed the items measuring these constructs could all be loaded into one composite score for overall teacher perception.

The second research question addressed in this portion of the study was: what are the factors related to teacher perceptions of new evaluation systems? The belief that the teacher evaluation system will increase teacher professionalism was significantly positively related to overall teacher perception of the evaluation system. Teacher professionalism was also the strongest predictor of teacher perception, where a one-unit increase in the teacher professionalism composite score indicates a 0.83 increase in the overall teacher perception composite score. This increase is nearly an order of magnitude larger than the impact of the next largest predictor.

Additionally, the factor 'Beliefs about the evaluator' was a slightly positive predictor of overall teacher perception. A one-unit increase in this factor indicated a 0.09 increase in the

overall teacher perception score. Thus, teachers with positive beliefs about their evaluators are more likely to have an overall positive perception of the teacher evaluation system, which existing literature supports. Finally, teacher familiarity with the teacher evaluation system is also significantly positively related to overall teacher perceptions of the system, showing that the more teachers feel they know about the evaluation system, the more likely they are to have a positive perception of it. A one-unit increase in familiarity with the evaluation system indicates a 0.10 increase in the overall perception score.

Taken as a whole, these results reveal that that two key beliefs, how the evaluation system supports teacher professionalism and beliefs about the evaluator, combine with perceived familiarity with the evaluation system, form the majority of the teacher's overall perception of the evaluation system. The next phase of the study further explored the how and why underpinning these relationships.

C. **Qualitative Findings**

The interview portion of the study was designed to gain a better understanding of why certain factors, as revealed in the survey portion of the study, were significantly related to teacher perceptions of the local evaluation system. First, while overall teacher perception of the local evaluation system was neutral, interviews revealed that while teachers largely felt positive around the design of the local evaluation system, including favorable opinions of the Danielson Framework for Teaching, they had concerns about the implementation of the local evaluation system, including the possibility of local evaluation system being used for retribution and concerns about evaluator objectivity. Second, in the interviews, teachers who had a positive perception of the local evaluation system spoke about the collaborative nature of their

conversations, and the opportunities they had to develop their goals and discuss feedback with their evaluators. The survey results showed that beliefs about the evaluator were significantly related to overall teacher perception of the local evaluation system. However, many teachers also expressed concerns about their evaluators, which centered on concerns about the objectivity of the evaluators and the potential for evaluators to develop a punitive culture around evaluation.

Additionally, the survey results showed a significant relationship between the teacher's perceived knowledge about the local evaluation system and their overall perception of the local evaluation system. The interviews provided additional detail here, revealing that the more teachers thought they knew about the Danielson Framework for Teaching, the more they liked the local evaluation system, and that most teachers felt they had received enough information and training about the local evaluation system. However, many teachers who had a positive perception of the local evaluation system said that they had heard stories and rumors from their colleagues about horrible experiences teachers in other schools were having with the local evaluation system, which led to many of those teachers offering various caveats to their favorable opinions.

Finally, the local evaluation system's perceived ability to boost the professionalism of the teacher workforce in River District was significantly related to an overall positive perception of the local evaluation system. Teacher interviews revealed that teachers felt that the local evaluation system, as originally designed in River District, provided a way to support reflective practice and teacher growth towards differentiated goals, and ultimately served as a way to move teaching in River District closer to the conception of teaching as a profession. However, the advent of PERA and Senate Bill 7 was seen as undermining the goals of the local evaluation system, and transforming the process from one that supported teachers as professionals to one

86

that used simplistic outcomes to determine which teachers would be affected by potential personnel actions.

1. **<u>Overall Perception</u>**

Overall teacher perception of the local evaluation system, as shown by the survey results, was neutral. However, subsequent interviews with teachers in River District showed that while the overall perception was neutral, most teachers had favorable impressions of the design and intentions behind the local evaluation system, but had misgivings about the implementation of the local evaluation system. Of the ten teachers who participated in the interviews, five of them had overall perceptions of the local evaluation system that can be described as positive, four teachers had mixed perceptions, and one teacher had a strongly negative perception of the local evaluation system. The comments from each group revealed common themes within the groups. Teachers who had positive perceptions of the local evaluation system all spoke about the way the local evaluation system supported their professional growth by encouraging reflection on their practice and conversations about instruction with their evaluators. The group of teachers with mixed opinions praised the design of the local evaluation system, but expressed concerns about the fidelity of the evaluators' implementation of the local evaluation system. The teacher with negative perceptions condemned the design, intentions, and implementation of the local evaluation system.

The teachers with positive overall perceptions of the local evaluation system praised the way the local evaluation system supported their professional growth and encouraged teachers to take ownership of their evaluation goals. One new high school teacher, who came to teaching after a career change, described the local evaluation system as "a very affirming process", and one that reinforced her desire to be a teacher, because she found the observation and conference

cycle to be encouraging. Additionally, she felt that the local evaluation system allowed her evaluator to give her specific, actionable feedback on her teaching, and allowed her to work with her evaluator in a non-threatening manner. Another high school teacher emphasized the way the local evaluation system acknowledged the professional nature of teaching, stating "most peoples' goals are to help us as professionals and treat us like professionals". Additionally, this teacher praised the design of the local evaluation system, stating

> I think that as it's created it's awesome. I think it's designed to help teachers become better teachers and to help us reflect on what we've done in our classrooms. And if it's done right, I think it could be really awesome for us, helping us grow as teachers.

A National Board Certified elementary school teacher also praised the design of the local evaluation system, citing the way it encourages reflection and forces teachers to really involve their supervisor. Finally, a teacher who was closely involved with the development of the local evaluation system said that she felt that the local evaluation system was fundamentally the right direction for her district.

The group of teachers with mixed perceptions of the local evaluation system praised many of the same elements cited by the teachers with positive perceptions. However, the teachers with mixed opinions expressed concerns about the fidelity of implementation as well as the way teachers might be held accountable using the results of the local evaluation system. For example, a special education teacher at a school that serves a large population of low-income students and English language learners stated:

> I'm a little nervous about [the local evaluation system]. I think it may be well-intentioned but I am very concerned, especially in a school with the at-risk populations such as this one has, that teachers are going to be held responsible for improvement they have no control over, which is essentially everything that happens outside the classroom. So many of these kids have factors that prohibit them learning to their fullest extent, and teachers, there's only so much control we have over that.

Several teachers in this group mentioned that they had negative experiences with evaluation under the previous system, usually with respect to an evaluator who was vindictive or extremely subjective in their assessment. For example, a reading specialist praised the local evaluation system as a "more teacher-oriented, teacher-friendly type of evaluation, through which there can be conversations that would not have been had with prior evaluations", but also stated that the local evaluation system can still be very subjective. Likewise, a veteran high school English teacher described the Danielson Framework for Teaching, which is the observational rubric for the local evaluation system, as a "brilliant idea" that addressed all important aspects of teaching, but then expressed skepticism about whether evaluators had enough time to implement the system with fidelity, rather than using it as a checklist.

Finally, one teacher had a strongly negative perception of the local evaluation system, even though she had received the highest possible rating during her evaluation. When asked her overall opinion of the local evaluation system, she described it as "dehumanizing and demeaning". She argued that it is "overly simplistic, even though on the face of it it's very complex and very well laid out". When asked to elaborate, this teacher described the local evaluation system as just a checklist designed to "quantify teaching". Furthermore, she argues that the evaluator is so constrained by the framework, that he or she misses the "wholeness" of what is happening in the classroom.

2.   **Evaluators, Culture, and Sensemaking**

Existing literature on teacher evaluation policy points to the centrality of the evaluator in teacher perceptions of those policies (McLaughlin, 1990; Stiggins & Duke, 1988; Milanowski & Heneman, 2001), which was supported in the survey results. Subsequent interviews with teachers

provided additional insight into how beliefs about the evaluator affected their opinion of the

overall system. When asked about how their relationship with the evaluator impacted their

perception of the local evaluation system, the participants' responses showed that most teachers

spoke highly of the local evaluation system's potential to improve instructional conversations

and provide more targeted feedback, especially when the teacher and the evaluator approached

the local evaluation system from a shared understanding of its design and goals. However, many

teachers expressed serious concerns about the objectivity of the evaluators and the potential for

the system to be used in a vindictive manner, or for a punitive culture of evaluation to develop.

Overall, teachers who had positive overall perceptions of the local evaluation system

spoke highly of the system's ability to give targeted, actionable feedback. One participant, a third

year high school teacher who has struggled with giving clear directions, described how her

evaluator has supported her in thinking through her lesson design to prevent student confusion.

She states, "'[my evaluator] has definitely given me very specific things that she's going to look

for on my informal observations and that's been good because I am trying to practice just a little

more time thinking through [the lesson]". Likewise, another high school teacher also praised the

feedback she had received from her evaluator, stating:

> When I first came to the district I had an evaluator that was very good at identifying what
> I needed to work on. She had an excellent way of communicating that without making me
> feel bad, because it was like my third year teaching and I still don't know what I'm doing
> sometimes.

Overall, teachers who had favorable overall impressions of the local evaluation system spoke

highly of the system's ability to encourage specific, actionable feedback regarding instruction.

Additionally, teachers with positive overall impressions described how they shared with

their evaluator an understanding of the purpose and design of the local evaluation system, and

how this shared understanding led to a more collaborative and useful evaluation. Many of these

teachers spoke about how they advocated for an open-door policy and encouraged their evaluators to visit their classrooms at any time. The teachers who spoke about this repeatedly mentioned they trusted their evaluator because that person viewed the local evaluation system as a tool to help teachers improve, not as a way to penalize teachers. For example, one elementary teacher with a favorable perception of the local evaluation system said she requested that her observation take place during the guided reading part of her day, since she felt that was the weakest part of her instruction. In describing her reasoning for what could have been a risky choice, she said:

> [That choice] comes back to the relationship piece because I knew it was about my professional growth and I trusted this person to give me honest feedback and that's what it was about for me. It was about growing my practice and I don't think if I didn't trust my principal I don't think I would have asked them to do that. That doesn't mean I would have dog and pony showed it, but I don't think I would have actively sought out a moment in my day where I felt weak.

Another teacher with a positive overall perception of the local evaluation system described how her evaluator, in this case her principal, built a school culture that treated evaluation as an opportunity for growth and helpful feedback. Regarding the culture around evaluation, this teacher said:

> I feel like that vision casting from the top was especially strong here. She really would come if you said "I'm trying something new if you have a chance stop by", she would and she brought a great perspective.

All of the teachers with positive overall perceptions of the local evaluation system made similar comments about how they found their evaluator to be accessible and trustworthy. Because the teacher and their evaluators shared an understanding regarding the goal of the local evaluation system – professional growth – the evaluative process became one in which those teachers felt comfortable revealing weakness and actively seeking feedback. However, since the evaluator's ability to develop a culture like this, as well as to provide opportunities for teachers to build

shared understandings with the evaluators, is related to overall perception, when an evaluator

fails to do so it can undermine the entire evaluative process. One teacher noted this in her

comments, stating:

> I have an administrator, a supervisor, who I feel respects me as a professional and it's a
> safe situation to go in and do that. I'm not quite sure that exists, that the culture exists, in
> many of our buildings.

In sum, the way in which the evaluator and teacher develop a shared understanding of the design

and goals of the local evaluation system, and the extent to which the evaluator cultivates a

collaborative, low-stakes relationship with the teacher can lead to more positive overall

perceptions of the process.

Regardless of what they believed about their evaluator, objectivity on the part of the

evaluator and the potential for abuse in the system was a concern for almost all teachers.

Perceived objectivity on the part of the evaluator was the first concern. Despite the use of the

Danielson Framework for Teaching as the evaluation rubric, teachers expressed concern that

evaluators would not be fair or objective in their appraisals. When pressed to define what he

meant by objective, one teacher described how in his most recent evaluation his evaluator was

able to tell him the exact words the teacher used with his class. By grounding her evaluation in

the precise instruction and events in the classroom, this teacher's evaluator reassured him that his

evaluation would be objective, and not affected by personal opinion. He stated that his evaluator

only discussed what she was able to observe in the classroom, which was his definition of

objective. Another teacher defined an objective evaluation as one that only used test scores, since

she thought that observational data would always be subjective. A high school English teacher

likened teacher evaluation to grading, stating, "I feel like there's a loophole no matter how you

structure [the evaluation], just like grading. No matter how objective you try to make it, it's

always going to be a little subjective". Finally, one elementary teacher with a positive perception

of the local evaluation system still expressed concern about the extent to which subjectivity plays

a role, and how it could undermine the consistency of the evaluation scores between schools.

This teacher states:

> I think that some teachers get concerned because it can be subjective you know. One
> principal's two can be another's three. Some principals, absolutely, no matter, will not
> give a four in their building because they feel that nobody would reach that, which I can
> understand. So that gets frustrating for teachers that maybe at school A I'd get a two but
> at school B I'd get a three.

Despite the robust observation rubric, even teachers who had positive views of the local

evaluation system were concerned that their evaluations would be affected by the evaluator's

feelings or personal judgments.

Interview participants mentioned this concern about objectivity in concert with concerns

about how the evaluation system may be abused and used for retribution or vindictive purposes.

Participants with neutral or negative overall perceptions of the local evaluation system were

more likely to express concerns that their evaluations could be, or are currently being used, to

punish certain teachers, usually for personal reasons. A veteran elementary teacher argued that

the personal feelings of the evaluator or teacher will affect the evaluation, stating:

> the other part of the subjectivity that concerns me is that you may have an administrator
> who simply doesn't like a teacher for some reason or even vice versus…people know
> who the administrators like and who they don't, who they trust and who they don't, or at
> least they perceive they do, perhaps that's a better way to say it. So there is all of those
> things that come into play.

One veteran high school teacher went further, stating, "there are teachers who are being

scapegoated because they are outspoken. So that's why I worry about the subjectivity of the

evaluation and we need to make sure it is objective and reliable across the board". This same

teacher also said that she felt teachers were less likely to assume teacher-leader roles because

they were concerned they would come under scrutiny and be penalized for being outspoken.

Another high school teacher made an even stronger statement, stating, "People are afraid.

Several of the [evaluators] use it as retribution". The interview participant who had a negative

overall perception of the local evaluation system believed that teachers who largely agreed with

the evaluators received the highest ratings. She said:

> I've noticed in my school that the people who have gotten the highest ratings, because word gets out, are the people who are the most sycophantic. I went woah, well there's a link there to me. It comes in anywhere. I think would come in no matter how you laid it out.

One difference that emerged between teachers with neutral or negative perceptions of the local

evaluation system and teachers with overall positive perceptions of the local evaluation system

was that teachers with positive perceptions, while still expressing concerns about retribution,

were careful to articulate that they did not believe this was true of most evaluators. Reflecting

this view, one middle-school teacher stated

> I think there are two sets of people who are involved with it. I think there are people who are using it to document for certain people, and I think the vast majority, I do think they want to improve the quality of instruction, the communication process between evaluator or appraiser and instructor. But I do think that it can be used in some ways in a punitive fashion.

A new high school teacher expressed a similar view, stating, "I'm grateful to have a good

working relationship here but I think the whole thing hinges on that and I do think that it can be

abused". Overall, almost all interview participants, even those with positive views of the local

evaluation system, felt that the system either could be, or was currently being, abused by some

evaluators.

3. **Knowledge about the Evaluation System**

In the survey data, teachers' perceived familiarity with the evaluation system was positively correlated with their overall perception of the system. In interviews, teachers elaborated on this relationship, revealing that teachers largely equated knowing about the evaluation system with knowing about the Danielson Framework for Teaching. The more teachers felt they knew about the Framework, the more likely they were to have a positive view of the overall evaluation system. This is because in all but one case, teachers who felt they really understood the Danielson Framework for Teaching thought the Framework reflected quality teaching and was an appropriate instrument to use in teacher evaluation. For example, one highly rated elementary teacher talked about how she went through the Framework row by row to understand the differences between the different scores. She determined that the difference between a four, the highest score, and a three was the extent to which the classroom was student-centered. Because this aligned with her own definition of good teaching, this teacher felt strongly that the evaluation system would support her professional growth and was a valid way of assessing teachers.

Six out of the ten interview participants explicitly praised the Danielson Framework for Teaching. One teacher who was closely involved with the design of the teacher evaluation system described how the Framework is seen as a district-wide definition of quality teaching, stating, "the Framework for Teaching is our agreed upon language and what we believe to be good teaching and it is that starting point for us to be having conversations around". A veteran high school teacher said the Framework addresses all of the important aspects of teaching, from planning to execution. Another high school teacher praised the way in which the Framework could be used to set teacher-specific goals. She stated, "There are a lot of things I really like

about Danielson because it lets you work at your own level, and that's ideally what I would like to see because everybody is in a different place". Finally, a middle school reading specialist described the Framework fostering more in-depth conversations about instruction.

While there was general support for the district's decision to build the teacher evaluation system around the Danielson Framework for Teaching, the interview participants disagreed about whether they had received adequate training and information about the evaluation process, and specifically the Danielson Framework for Teaching. Teachers who said that they felt they had received enough information about the evaluation process and the Framework usually had been exposed to the Framework in other ways, usually through the district's new teacher mentoring process either as a mentor or a mentee. Additionally, interview participants who were school representatives for the local teachers union or involved in the union in other ways reported that they received ample information from the union, and felt that they had a better understanding of the overall process. One teacher said, "I'm part of the union so I understand the concept behind the development of [the evaluation system] and I think that's probably why I have a more positive outlook than some people do". Finally, all of the National Board Certified interview participants described how the reflection prompted by their certification process was similar to the evaluation cycles in the teacher evaluation system.

Other teachers felt strongly that they had not received enough information about the evaluation process, despite a district-wide professional development initiative targeted at training teachers. During the interview, one special education teacher at a school with a large number of bilingual programs said the staff at her school had been discussing scheduling another professional development to address lingering confusion about the evaluation process. Similarly,

another teacher said that teachers in her school were still confused about the evaluation process.

She stated:

> And [the district] really gave you a lot of professional development on it, and you know people were going through a lot of professional development, we were talking a lot about it at our staff meetings, we were getting a lot of emails from the union and administration. A lot of teachers are still very confused because there is a lot. There is a lot of forms, there is a lot of cycles.

Another high school teacher made a similar comment, asserting that teachers had not received enough professional development to begin with, and that there should have been more of an effort to streamline it across the district. A veteran elementary teacher said she felt that teachers themselves were not taking the initiative to read the Framework and accompanying evaluation documents.

In addition to the perception that teachers did not receive enough professional development and training on the evaluation process, other interview participants argued that the evaluators did not receive enough training. One teacher who was involved with the development of the teacher evaluation system described how the initial evaluator trainings were effective, but the professional development for later evaluators was not as high quality. She said:

> We don't have enough opportunities to be doing trainings for our new administrators and so there is a disconnect around that. They're not doing anything wrong, they're following the training they went to and they're not getting enough information back out of our system to tell them where we are.

Similarly, a middle school teacher expressed skepticism about whether evaluators were adequately prepared to evaluate different teaching styles, stating:

> I'm not sure that they've had enough development in terms of different types of teachers and what to look for. My evaluator told me they give them a checklist in training. So how different administrators look at these checklists worries me a little bit, because they have unintentional biases that they might not be aware of.

Finally, a veteran high school English teacher spoke about what she perceived to be an imbalance between the way teachers were held accountable under the evaluation system and the extent to which evaluators were held accountable for knowing how to evaluate teachers in a fair way that was faithful to the spirit of the evaluation system.

In addition to the district trainings and ongoing professional development, many teachers also received some or almost all of their information about the evaluation system from their colleagues, especially information that emphasizes potential negative consequences from the new evaluation system. Furthermore, they also heard stories of other teachers' experiences with the system. Six out of the ten interview participants stated that even though they were largely happy with the evaluation system, they had heard at least one story about another teacher's bad experience with the system. Almost every teacher with a positive overall perception brought up that they had heard horror stories from colleagues, even though they were generally happy with the system. Some teachers alluded to schools where the culture of evaluation was punitive. One teacher, in discussing her experiences with evaluation, remarked, "We– and I mean this pretty cross-district -- districts know the sites where the culture is worse, people know within in the district. People know where the issues are, people know where the cultures are bad". Another high school teacher simply states, "people are afraid", while an elementary teacher says, "people are worried about this. They are conversations going on all over the place". Even a new teacher who spoke highly of her experiences with the evaluation process thinks that the process could be abused and told a story about a colleague at another school in the district as an example of such abuse.

In one sense, these stories and rumors provide a different, though perhaps no less valuable, channel for information than the formal trainings and professional developments

provided by the district. While it is true that the more teachers felt they knew about the teacher

evaluation program, especially the Danielson Framework for Teaching, the more likely they

were to have a high opinion of that program, even teachers with positive overall perceptions had

heard stories of the evaluation system gone awry.

4.   **Teacher Evaluation at the State and Local Levels**

Almost all interview participants expressed some concern and frustration with the

adoption of the state requirements. Most teachers felt that the state requirements undermined

their local evaluation system, reducing teacher evaluation from an opportunity for feedback and

growth to something teachers had to endure in order to keep their job. One teacher described the

adaption of the local district system as follows:

> The local system is professional practice and growing professional practice and
> professional development. The state requirements are accountability and quality
> assurance, bottom line. They don't intersect. And that's why we made the choice we
> made, in how we layered it on, because, um, [pause], we philosophically needed to find a
> way, I don't want to say circumvent the system but to make sure that we met the
> requirements of the law but still maintained our very strong beliefs.

Another teacher described the process using stronger language, stating, "My true words on that,

they've bastardized it", referring to the local evaluation system. She also said:

> the state had to step in and ruin the whole process, to be honest with you, by ranking us
> and putting us in buckets…it's totally taken away from the integrity of the program
> because the state interfered with what was a good program to begin with.

Finally, this same teacher also expressed frustration that the state legislators did not look, at least

to her knowledge, to districts with strong existing evaluation programs, such as River District, as

a model for the state's requirements.

When asked about how each evaluation score is used, all interview participants

responded that the local evaluation system's score is used for professional growth and

development while the state evaluation score is used to determine who is laid off, if the district

has to make cuts. A high school special education teacher said:

> The local system, and I may be wrong, the local system results are how are you doing in the classroom, how proficient are you, are the students learning, are they actively engaged, are you able to communicate clearly and accurately to the students. The local score then feeds into the state score, and that you go on a wonderful list on whether you will be on the chopping block.

Another teacher made a similar remark, "It's insane. It's pretty confusing, but that's pretty much

what [the local score] being used for, to determine your state rating. Originally, it was used to

help you decide where you needed to improve and where you were at". It clearly emerged from

the interviews that the teachers in River District felt that the purpose and goals of their existing

evaluation system had been completely undermined by the state teacher evaluation requirements.

The findings in the interview portion of the study made it clear why teachers felt the

professionalism of the teacher evaluation system and their beliefs about their evaluator, as well

as perceived familiarity with the system, were significant factors in overall perception. While the

survey results show that overall teacher perception of the evaluation system is neutral, interviews

show that teachers have positive opinions of parts of the evaluation system, especially the

Danielson Framework for Teaching. Moreover, teachers who feel knowledgeable about the

Danielson Framework for Teaching genuinely see it as a vehicle for improved instructional

conversations and feedback. However, teachers also have serious misgivings about the system's

potential for abuse, including the use of evaluation as retribution, and are frustrated about the

implementation of the state evaluation requirements and how those requirements have altered the

local evaluation system. Taken together, these findings reveal that the extent to which teachers

see the local evaluation system as moving towards an evaluation system built around a

conception of teaching as a profession, including collaborative conversations with their evaluator

and opportunities for growth, have a more favorable view of the system. However, the state

teacher evaluation requirements are undermining this movement towards professionalism, which

has implications for the overall success of the local evaluation system in River District, which

will be addressed in the chapter.

VI. ANALYSIS

A.   **<u>Introduction</u>**

The purpose of this study was to examine the factors that influence teacher perceptions of new accountability-oriented teacher evaluation policies, such as the one currently being implemented in River District. The combination of survey and interview results in this study allowed for the identification of factors that were significantly related to teacher perceptions of evaluation policies and provided insight into how these factors are related to overall perception and to each other. An examination of these findings shows that each level of the cognitive implementation framework – policy signals, organizations, and individual – a significant factor could be found. Additionally, teacher perceptions of the new evaluation policies cannot be understood by examining factors in isolation from one another. Understanding the interaction of the factors within and across the levels is necessary for understanding how teacher perceptions of the new evaluation system.

Overall teacher perceptions were significantly related to the extent to which the system, especially the Danielson Framework for Teaching, was seen as supporting the professionalism of the teacher workforce. The perceived support for professionalism was a policy signal to which teachers responded. Additionally, at the organizational level the evaluator, especially the extent to which teachers felt they shared with their evaluator an understanding of the purpose of the evaluation, played a significant role in teacher perceptions. Finally, at the individual level, teachers' perceived familiarity with the evaluation system, including the formal and informal ways teachers learned about the evaluation system, was also a significant predictor of overall perception.

The analytical approach to this study was grounded in a cognitive implementation framework (Spillane, Reiser, & Reimer, 2002), which suggested that factors on the policy level, organizational level, and individual level would all play a role in how teachers perceived a new evaluation policy. The identified factors, as well as teachers' explanations of how and why each factor was related to overall teacher perceptions, were analyzed at each level of the cognitive implementation framework. This approach was in contrast to a simplistic explanation that relied only upon one level of analysis to explore why teachers view evaluation policies in certain ways, and proved to be the case in this study. At each level of analysis, a significant factor was identified. At the policy signals level, the perceived professionalism of the evaluation system was a significant factor. At the organizational level, teacher beliefs about the evaluator were a significant factor. These levels also affected one another. Evaluators, at least reported by the teachers, faced pressure to implement the state policies in way that directly undermined the professional nature of the evaluation system. At the individual level, teacher's perceived familiarity with the evaluation system was a significant factor. The organizational and individual levels also influenced one another. Teachers reported the most positive beliefs about the evaluation process when they worked from a shared understanding with their evaluator of the evaluation process. This shared understanding hinged partially on what teachers believe they know about the evaluation system.

Taken together, the factors that influence overall teacher perception of evaluation policies, and the relationship between those factors and overall perception, highlight a tension in River District's teacher evaluation system between a professional approach to teacher evaluation and a bureaucratic approach to teacher evaluation. At all three analysis levels – policy, organizational, and individual – this tension is evident. In fact, the teacher evaluation system in

River District was explicitly designed to facilitate teacher improvement through a collaborative process that emphasized opportunities for growth and hewed closely to Darling-Hammond's (1990) argument that the most effective, growth-oriented evaluation systems emphasize "personal goal setting and self-evaluation, joined with peer-mediated and situationally relevant reviews of practice" (39). However, the changes made to the evaluation system to comply with PERA and SB 7 have pushed River District towards a more bureaucratic approach to teacher evaluation, and therefore moved the district away from the desired type of evaluation. The bureaucratic approach to evaluation emphasizes standardization, treats all teachers alike, and values compliance with procedures (Darling-Hammond, 1990).

In River District, teachers saw the state requirements as pushing the district towards a more bureaucratic approach to evaluation, which diminished their support of their local system. However, River District is already had a strong local system in place. If the local evaluation was nonexistent, perfunctory, or unfair, a push towards a bureaucratic approach could improve the system by providing an increased level of fairness and standardization, while a professional approach might lie outside of that district's capacity at that time. The districts that are most likely to benefit from a state level mandate, even if it takes a bureaucratic approach to evaluation, are the districts that are struggling with teacher evaluation. The districts with strong evaluation systems are likely to see that work undermined by a state level mandate. A bureaucratic approach will help struggling districts, while a professional orientation will support the work that is already happening in leading districts. These approaches are in tension with each other, since they are grounded in fundamentally different conceptions of teaching. On one hand, the professional approach to teacher evaluation is more likely to lead to instructional improvement, but this may be too ambitious for struggling districts. On the other hand, a bureaucratic approach

will provide some minimal level of quality, but may actually dissuaded teachers in leading districts from engaging with the policy in the desired way. The remainder of this chapter will discuss how this tension emerges at each level of analysis in turn, as well as links between the levels.

B.   **Policy Signals Level**

At the policy signals level, an analysis of the study results shows that the changes made to the River District evaluation system in order to comply with PERA and SB 7 will make it less likely that the teacher evaluation system will drive instructional improvement in the district. This is because the changes made to the evaluation system and how the evaluations are used shifted the district's approach to evaluation from a perceived professional orientation to a perceived bureaucratic orientation. Accordingly, teachers will be less likely to engage with the evaluation system in a meaningful way, since the perception that the teacher evaluation system will increase the professionalism of the teaching workforce is the factor most strongly related to overall positive perception. Furthermore, teachers spoke at length about their beliefs that the River District evaluation system treated the teachers as professionals, and that the state requirements did not share this approach. Teachers also conveyed the sense that the goals and purpose of the teacher evaluation system in River District were undermined by PERA and SB7.

River District already had a complex teacher evaluation system, which teachers perceived as built around a conception of teaching as a profession (Wise et al., 1984) that allowed the district to adopt a professional orientation towards teacher evaluation. Indeed, professionalism, trust, collaboration, and continuous learning are all identified as core values of the teacher evaluation system in River District. This orientation can be seen throughout the design of the

evaluation system, and aligns with Wise and Darling-Hammond's (1984) description of a

professional evaluation system. To begin, the district chose to use the word 'appraisal' rather

than 'evaluation' to signal that the process was designed to assess opportunities for growth.

Teachers were required to develop their own evaluation goals, which were customizable to a

particular teacher's situation, and to develop a professional development plan to assist them in

meeting those goals. Additionally, the district chose to operate from a presumption of

competence, rather than use the evaluation process as a way to monitor for minimal performance

levels. Finally, the district built the evaluation system around the Danielson Framework for

Teaching, which became River District's chosen definition of effective teaching.

In contrast, teachers viewed the state requirements are built around a conception of

teaching as labor (Wise et al., 1984), in which outputs, namely student test scores, are seen as

measures of quality, and teaching is seem as something that can be standardized. This aligns with

a bureaucratic approach to evaluation, described by Darling-Hammond (1990) as follows:

> Bureaucratic evaluation relies on administrators (chiefly principals) to assess teachers in
> a standard manner using general criteria, such as generic teaching skills (does the teacher
> plan? set objectives? teach to the objectives? cover the curriculum?) or other context-free
> teaching behaviors. Bureaucratic evaluation is highly standardized, procedurally oriented
> and organized by checklist. (38)

Though teachers perceived the state requirements as a radical departure from the existing

evaluation system, this was not the case The state requirements actually led to only one major

change in the existing teacher evaluation system – linking the result of the evaluation to

personnel decisions. While River District did have to make some changes in response to the state

requirements that moved their approach closer to a bureaucratic approach, such as incorporation

of student performance, as measured by state standardized tests, the majority of the local system

remained unchanged. The local evaluation scores reflected the custom goals teachers developed

106

for themselves in collaboration with their evaluator, and the state evaluation scores were simply conversions from the local scores.

Teachers viewed the state and local teacher evaluation systems as having different approaches to evaluation, despite the fact that the bulk of the existing evaluation system remained unchanged. This is because, while the designs of the evaluation systems were fairly similar, the new state requirements changed the policy signals from the evaluation systems. By requiring that teacher evaluation scores be used for personnel decisions, the state requirements radically shifted the message about teachers and teaching conveyed by the policy, which in turn affected how teachers perceived the evaluation system (Spillane, Reiser, & Reimer, 2002). Ultimately, teachers, at least in River District, perceived the state requirements as undermining the local evaluation system, and transforming the evaluative process from one that promoted growth to an exercise to be endured in order to stay employed. This was the case in River District, which had a well-developed teacher evaluation system, but might have not been the case in another district with a haphazard or under-developed teacher evaluation system.

Attaching personnel decisions to teacher evaluations had several effects. First, the perceived purpose of the evaluation system was altered from one of growth and professional development to one of accountability, which was likely to change how teachers engaged with the system (Duke & Stiggins, 1990; Kimball, 2002). Teachers made a clear distinction between the purpose of their local evaluation score and their state evaluation score, but often noted that the local score was now essentially used to determine the state score, which made the purpose of both systems accountability. Second, by changing the purpose of the evaluation system, the state requirements shifted the image of River District's teachers from assets who could be developed to liabilities that needed to be managed and removed if necessary. Adler & Borys, (1996) in their

study of organizational design processes, describe this as the difference between organizational members who are seen as the source of the problems and organizational members who are seen as a source of "skill and intelligence". This is certainly not the case in all districts. If the local system already treated teachers as liabilities, then a bureaucratic approach might provide some welcome clarity and standardization. In River District, the evaluation process was now seen as producing scores designed to label, and if needed, remove, ineffective teachers rather than identifying and helping struggling teachers. These changes in the messages conveyed by the evaluation system changed how teachers responded to the system, with many teachers concluding that the existing evaluation system's commitment to teacher professionalism had been undermined by the state requirements.

C.   **Organizational Level**

These new state level policies also created pressure to adopt a bureaucratic approach to evaluation at the organizational level. To combat this pressure, teachers reported that their evaluators, usually principals, adopted a buffering role between the local evaluation process and the state requirements. Teachers who supported the evaluation system reported that their evaluators approached the evaluation process according to a shared understanding of a professional orientation towards teacher evaluation, regardless of what the state mandated. Additionally, teachers with favorable opinions of the evaluation system spoke about how their evaluators developed a culture of growth around evaluation, as opposed to the punitive culture many teachers feared. In describing how their evaluators made the evaluation process meaningful, useful, and safe by working from a shared understanding about evaluation, the teachers were describing how their evaluators engaged with the external demands of the state

108

policy, but in strategic ways. This aligns with Honig and Hatch's (2004) definition of buffering. The evaluators did not have the power to ignore the state evaluation requirements. Certainly, the evaluators did not give every teacher a high rating to ensure that none of their teachers were affected by any potential personnel actions. Rather, the evaluators minimized the impact of the new state requirements by making sure that teachers felt safe being honest and seeking help during the evaluation process. In doing so, evaluators emphasized the professional aspects of the evaluation system by making the focus of the process instructional improvement.

The role of evaluator as buffer also draws attention to the role of culture in teacher perceptions of the evaluation system. Concomitant with the formation of shared understandings is the evaluator's ability to build a culture of evaluation that focused on how evaluation could be used to support professional growth. In the interviews, almost all teachers expressed concern about the role of retribution or a punitive culture around evaluation. This captures the dual role culture plays in how teachers view the evaluation system. On the one hand, a culture that takes a professional approach to teacher evaluation is vital to successful evaluation (Iwanicki & Rindone, 1995; McLaughlin & Pfeiffer, 1988). On the other hand, given the centrality of culture to teacher perceptions of evaluation systems, if a punitive or retributive culture of evaluation develops it could have a disastrous impact on teacher perceptions of the evaluation system, especially since most teachers already have a sense that the state requirements are moving away from a professional orientation towards evaluation.

Overall, teachers see the state requirements as promoting a more bureaucratic approach to evaluation. Because of this, evaluators in River District are forced to use sensemaking and developing positive cultures around evaluation in order to address these concerns, and to encourage teachers to engage with the evaluation system in a meaningful way. In doing so,

evaluators in River District act as buffers between the state and local evaluation policy demands. In turn, this makes it more likely that teachers will perceive the evaluation system in positive ways. However, the centrality of the evaluator in this process also creates the possibility for punitive cultures around evaluation, especially since the state requirements have raised the stakes attached to the evaluation. In short, based on the findings from this study, teachers hope for the best when it comes to evaluation, but prepare for the worst.

D.   **Individual Level**

At the individual level, results indicate that not only is a shared understanding of the purpose and method of evaluation crucial to positive teacher perceptions, teachers are more likely to engage with the evaluation system in desired ways when the system aligns with their beliefs about high quality teaching. The finding that perceived familiarity with the evaluation system is significantly, positively correlated with overall perception of the evaluation indicates that the more teachers think they know about the evaluation system, the more likely they are to support the system. Teachers also had generally high opinions of the Danielson Framework for Teaching, and thought it captured most important aspects of high quality teaching. Taken together, these findings suggest that the more teachers believe they know about that system, especially the Danielson Framework for Teaching, the more teachers see that the evaluation system is built around a definition of high-quality teaching that aligns with their own beliefs.

The cognitive implementation framework posited by Spillane, Reiser, and Reimer (2002) argues that at the individual sensemaking level, teachers developed their perceptions through their individual understandings and schemata about the target of the policy. In this case, the more teachers see that the local evaluation system is built around a conception of teaching as a

profession, the more likely they are to support it, since most teachers view themselves as professionals (Sachs, 2001). This is consistent with the other findings in this study as well. Likewise, when teachers see the state requirements as built around a conception of teaching as industrial labor, based on the emphasis on outputs, this usually fails to align with their definitions of quality teaching, which will make it more likely that teachers will respond to these reforms in a negative way.

Given this, it is crucial to think about the formal and informal ways individual teachers learn about the evaluation system. Providing ample formal opportunities, such as professional development, mentoring, or presentations from the teachers union, will increase the likelihood that teachers will understand the purpose and underlying conception of teaching embedded in the evaluation system. In River District, teachers felt that they did not receive enough professional development or training about the mechanics of how the evaluation system worked. Several teachers spoke about how, by taking the time to dig deeply into the Danielson Framework for Teaching either alone or as part of a district professional development, they saw strong alignment between their beliefs about teaching and the core component of the evaluation system.

The need for ample formal opportunities for learning about the evaluation system is unsurprising. However, the results from this study also highlight the role of informal opportunities for learning about the evaluation system. Many teachers in the study reported that their colleagues or friends had been a source of information about state requirements for teacher evaluation. These informal learning opportunities that occur through conversations with colleagues and friends highlight the role of a teacher as a "sensegiver" about the teacher evaluation system in River District. Louis, Mayrowetz, Murphy, and Smylie (2013) define sensegivers as "people in the setting who understand the change goals, the school's culture and

111

history, and who are capable of communicating scenarios of consistency to others" (p. 43). When teachers act as sensegivers regarding teacher evaluation in River District, they can help teachers to understand the extent to which there is alignment between the conception of teaching embedded in the evaluation and the teacher's beliefs about high-quality teaching.

While teachers acting as sensegivers can potentially improve teacher perceptions of the overall evaluation, this hinges on the sensegivers understanding and supporting the evaluation system. Many teachers in this study, even those who strongly supported the evaluation system, referenced negative experiences with evaluation they have heard about secondhand, or referenced vague rumors about buildings with toxic cultures regarding evaluation. The state requirements attaching evaluation scores to personnel decisions have only heightened the sense of fear and trepidation about the possibility of negative evaluation experiences. This observation points to the power of stories in shaping teacher perceptions of the evaluation system. In their work on the power of storytelling in organizations, Swap, Shields, and Abrams (2001) argue that stories are a more powerful and memorable way of conveying information than more formal channels. Given this, countering negative stories about the evaluation system only heightens the need for teachers to convey how the local evaluation does in fact support a vision of teaching that is grounded in supporting teachers as professionals. Without this, it is likely that the district will face an uphill battle in developing positive teacher perceptions of the evaluation system.

E.   **Implications**

Overall, the findings from this study draw attention to the tension between local and state level policy-making. In the case of River District, this tension is evident in the dueling orientations towards evaluation, and how state legislation is exerting a push toward a more

bureaucratic approach to teacher evaluation. At every level of analysis – policy, organization, and individual – this tension affects how teachers view the evaluation system. At the policy signals level, teachers see the use of evaluation scores in personnel decisions as completely undermining the goals of the local evaluation system and substituting a bureaucratic system, focused on checklists and standardization, in its place. At the organizational level, to manage this tension, evaluators act as buffers between the state requirements and the local evaluation system by emphasizing shared understandings and professional evaluation cultures. At the individual level, teachers acted as sensegivers and storytellers, providing an opportunity for teachers to learn about the evaluation system outside of formal channels, but also allowing for stories of negative experiences to percolate throughout the workforce.

These findings have several implications. First, these findings highlight how state and local policy-making efforts can clash, potentially undermining one another. Often, state-level policies focus on using different policy instruments to elicit desired actions or changes at the local level. However, insufficient attention is paid to how state-level policies interact with and can be affected by existing local policies, especially if these policies are in some way more sophisticated or effective than those mandated by the state. River District illustrates this by showing how a robust local policy can be changed, and not in a positive way, by a state-level policy. Given this, policymakers should be mindful of how state or federal level policies may impact districts that have more sophisticated teacher evaluation systems already in place, as these policies might undo the work that has been done at the local level. This is particularly troubling given recent federal efforts, especially the Race to the Top competition, to focus policy attention the leaders rather than the laggards.

However, River District is perhaps a rare case in that it had a teacher evaluation system that had a professional orientation, and was developed through a collaborative process that involved all stakeholders including the teachers union. If a district had a weak or haphazard evaluation system, the push from the state regulations may provide a welcome way to regulate and standardize the evaluation process. Additionally, while in this case shared understandings and school cultures served as buffer from the demands of the state system, those are also the very things can be difficult to target directly with policy. This is especially true when policies from different levels of government, in this case state and local government, are not working in concert.

Second, the findings from this study highlight the malleability of perceptions of policy. Teacher perceptions of new evaluation systems are crucial for the successful implementations of those systems, but these perceptions are not necessarily based upon facts. For example, in this study, teachers generally had favorable opinions of the design of the evaluation system, and spoke of it in favorable terms, but complained bitterly about how the state requirements have changed the local system. It is possible that teachers only spoke so highly of the existing evaluation system because they are so unhappy with the new state requirements for teacher evaluation. Furthermore, teachers are receiving a variety of messages and information about the changes required by the state, and these messages are undoubtedly affected by the context and political landscape surrounding evaluation in River District. For example, it might be to the district's benefit to portray the required changes in a negative way. The local teachers' union might want to drum up support for the local evaluation system, since they were involved in the creation of that system and had no input into the state requirements..

F.   **<u>Directions for Future Research</u>**

This study identified several factors that were significantly related to teacher perceptions of new accountability-oriented evaluation systems and used interviews to explore how and why those factors were related to overall perceptions. First, further research should continue to identify factors and seek to understand the relationship between those factors and overall perceptions. Furthermore, this study hinged entirely on perceptions, rather than changes in practice. Further studies should compare teacher perceptions of the evaluation system with actual changes in teacher practice. The role of content area in teacher perceptions should also be studied, to see if teacher beliefs about their content area, and content specific teaching practices, impacts how teachers respond to evaluation policies. Finally, while teachers spoke at length about their evaluators, further research could continue to explore how the evaluators perceived, and changed their practice, based on new teacher evaluation policies.

This study was built upon a cognitive implementation framework, which argues the individual, the organization, and the policy itself, as well as the interaction between those levels, affect the sensemaking process around a new policy or policy change. Further research on teacher evaluation, and how teachers view and respond to those policies, should continue to explore the sensemaking process, particularly the ways in which formal and informal opportunities for sensemaking influence teacher perceptions of these policies. In particular, the link between more time and opportunities for sensemaking and positive perceptions should be further explored, in order to see if there is some optimal amount and type of opportunities for sensemaking. It is possible that too much time for sensemaking could create an echo chamber effect, in which negative perceptions are continually reinforced and amplified, while too little

time could result in increased teacher resistance or only surface level compliance with the new policy.

Additionally, future research should examine the governance issues that emerge from the clash of state and local policies. In River District, these issues led to a general sense that the state level policy undermined the progress made in recent years at the local level. However, this is unlikely to be the case in other localities, and future research should continue to explore the different ways local and state policies change each other, and how the clash of those policies may support or undermine efforts to improve student learning.

G.   **Conclusion**

This research contributes to existing research on teacher evaluation in two ways. First, it joins the growing body of research on the national wave of teacher evaluation policies spurred on by the Race to the Top competition, but provides a unique perspective due to the suburban location of this study. Additionally, by using a cognitive implementation approach, this research considered the ways in which the policy, the organization, and individual characteristics all contributed to how teacher viewed these policies. In particular, the use of this framework allowed this research to explore not only the individual sensemaking process, but the situated sensemaking that takes place in an organization, and how informal and formal opportunities for sensemaking can influence that process. Second, this research provides a way to explore how a district with strong existing evaluation policies responds to a state mandate that is not always aligned with the goals of the district. The implementation issues raised by the clash of local and state policies draws attention to the ways in which top-down mandates can undermine or hinder local improvement efforts, even without requiring dramatic changes to existing policies.

116

In sum, teacher evaluation has been the subject of increased policy attention and action in recent years. These policies are designed to improve teaching and learning by improving the quality of the teacher workforce, both through increased feedback and support and targeted personnel actions. Teachers in River District, a large suburban district in Illinois, largely support their local teacher evaluation system because they see it as supportive of them as professionals, but have serious misgivings about the effect of recent legislation in Illinois regarding teacher evaluation. The changes made to the local system, namely requiring the use of evaluation scores in personnel actions, are seen by teachers as promoting a bureaucratic approach to teacher evaluation that is in tension with the professional approach embedded in their local evaluation system. This tension is evident at the policy, organizational, and individual levels. The evaluator and teacher respond to that tension by using sensemaking, culture, and storytelling to either buffer the external policy demands or disseminate information through informal channels. However, at best these are attempts to manage this tension. To truly see teachers engage with the evaluation system in a meaningful way and to achieve the goals of these reforms, teacher evaluation systems at all levels need to be built around a professional approach to teacher evaluation.

CITED LITERATURE

Adler, P. S., & Borys, B. (1996). Two types of bureaucracy: Enabling and coercive.
Administrative Science Quarterly, *41*, 61–89.

American Recovery and Reinvestment Act. Pub. L. 111-5 (2009).

Ashforth, B. E., & Mael. F. (1989). Social identity theory and the organization. *The Academy of Management Review, 14*(1), 20-39.

Bacharach, S. B., Conley, S. C., & Shedd, J. B. (1990). Evaluating teachers for career awards and merit pay. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 133-146). Newbury Park, CA: Sage Publications.

Brandt, R. M. (1995). Teacher evaluation for career ladder and incentive pay programs. In D. Duke (Ed.), *Teacher evaluation policy: From accountability to professional development.* (pp. 13-34). Albany, NY: State University of New York Press.

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis, 12*(3), 311-329.

Cohen, D. K., & Moffitt, S. L. (2009). *The ordeal of equality: Did federal regulation fix the schools?* Cambridge, MA: Harvard University Press.

Cohen, D. K., Moffitt, S. L., & Goldin, S. (2007). Policy and practice: The dilemma. *American Journal of Education, 113,* 515-548.

Converse, P.D., Wolfe, E.W., Huang, X., & Oswald, F. L. (2008). Response rates for mixed-model surveys using mail and e-mail/web. *American Journal of Evaluation, 29*(1), 99-107.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2011). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Upper Saddle River, NJ: Pearson Education.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, (pp. 209-240), Thousand Oaks, CA: Sage.

Darling-Hammond, L. (1990). Teacher professionalism: Why and how? In A. Lieberman (Ed.), *Schools as collaborative cultures: Creating the future now* (pp. 25–50). Bristol, PA: The Falmer Press.

Darling-Hammond, L. (2000). Teacher quality and student acheivement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1). Retrieved from *http://epaa.asu.edu/v8n1*.

DeBray, E. H. (2006). *Politics, ideology, and education: Federal policy during the Clinton and Bush administrations.* New York, NY: Teachers College Press.

DeBray-Pelot, E. H., & McGuinn, P. J. (2009). The new politics of education: Analyzing the federal education policy landscape in the post-NCLB era. *Educational Policy, 23*, 15-42.

Department of Education. (1991). *America 2000: An education strategy* (Report No. 0S91-13). Washington, DC: U.S. Government Printing Office.

Duke, D. L., & Stiggins, R. J. (1990). Beyond minimum competence: Evaluation for professional development. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing elementary and secondary school teachers.* (pp. 116-132). Newbury Park, CA: Sage Publications.

Elmore, R.F. (2002). Unwarranted intrusion. *Education Next, 2*(1), 31-35.

Enrich, P. (1995). Leaving equality behind: New directions in school finance reform. *Vanderbilt Law Review, 48,* 101-187.

Evans, R. (1996). *The human side of school change: Reform, resistance, and the real-life problems of innovation.* San Francisco: Jossey-Bass.

Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis*, *13*(2), 189-202.

Fuhrman, S. H., Goertz, M. E., & Weinbaum, E. H. (2007). Educational governance in the United States: Where are we? How did we get here? Why should we care? In S. H. Fuhrman, D. K. Cohen, & F. Mosher (Eds.) *The state of education policy research* (pp. 41-61). Mahway, NJ: Lawrence Erlbaum Associates.

Fullan, M. (2007). *The New Meaning of Educational Change* (4th ed.). New York: Teachers College Press.

Greene, J. C., & Caracelli, V. J. (2003). Making paradigmatic sense of mixed methods practice. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, (pp. 91-110), Thousand Oaks, CA: Sage.

Goals 2000: Educate America Act. Pub. L. 103-227 (1994).

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Washington, D.C.: National Comprehensive center for Teacher Quality.

Guadagnoli, E., & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin*, *103*(2), 265.

Hannaway, J., & Woodroffe, N. (2003). Policy instruments in education. *Review of Research in Education, 27*, 1-24.

Harrington, M. (1962). *The Other America: Poverty in the United States.* New York, NY: Macmillian.

Hill, P. T. (2001). *The federal role in education.* Washington, DC: Brookings Institution.

Hill, P. T. (2007). New political economy of public education: Policy and research. In S. H. Fuhrman, D. K. Cohen, & F. Mosher (Eds.) *The state of education policy research* (pp. 87-104). Mahway, NJ: Lawrence Erlbaum Associates.

Honig, M. I. (2006). Complexity and policy implementation: Challenges and opportunities for the field. In M. Honig (Ed.), *New directions for education policy implementation* (pp. 1-24). Albany, NY: State University of New York Press.

Illinois State Board of Education. (2011). *What new evaluations would mean for teachers.* Springfield, IL: Author.

Ingram, H. M., Schneider, A. L., & DeLeon, P. (2007). Social construction and policy design. In P. A. Sabatier (Ed.), *Theories of the Policy Process*, (pp. 93-126). Boulder, CO: Westview Press.

Iwanicki, E. F., & Rindone, D. A. (1995). Integrating professional development, teacher evaluation, and student learning: The evolution of teacher evaluation policy in Connecticut. In D. Duke (Ed.), *Teacher evaluation policy: From accountability to professional development.* (pp. 35-64). Albany, NY: State University of New York Press.

Jacob, R. T. (2011). An experiment to test the feasibility and quality of a web-based questionnaire of teachers. *Evaluation Review, 35*(1), 40-70.

Jansen, J.D. (2001). Image-ining teachers: Policy images and teacher identity in South African classrooms. *South African Journal of Education, 21*(4), 242.

Kingdon, J. W. (2011). *Agendas, alternatives, and public policies* (Revised 5th ed.). New York, NY: Longman.

Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal, 48,* 361-386.

Krosnick, J. A. (1999). Survey research. *Annual review of psychology*, *50*(1), 537-567.

Lewis, O. (1954). *Five families: Mexican case studies in the culture of poverty.* New York, NY: Basic Books.

Linder, S. H., & Peters, B. G. (1989). Instruments of government: Perceptions and contexts. *Journal of Public Policy*, *9*(01), 35-58.

Louis, K. S., Mayrowetz, D., Murphy, J., & Smylie, M. (2013). Making sense of distributed leadership: How secondary school educators look at job redesign. *International Journal of Education Research*, *1*(1), 33–68.

Manna, P. (2006). *School's in: Federalism and the national education agenda*. Washington, DC: Georgetown University Press.

Manna, P., & Petrilli, M. J. (2008). Double standard? "Scientifically based research" and the No Child Left Behind Act. In F. Hess (Ed.). *When research matters: How scholarship influences education policy* (pp. 63-88). Cambridge, MA: Harvard Education Press.

McDonnell , L. M. (2005). No Child Left Behind and the federal role in education: Evolution or revolution?. *Peabody Journal of Education, 80*(2), 19-38.

McDonnell, L., & Elmore, R. (1987). Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis, 9*(2), 133-152.

McGuinn, P. J. (2006). *No child left behind and the transformation of federal education policy, 1965-2005.* Lawrence, KS: University Press of Kansas.

McGuinn, P. (2010). *Creating cover and constructing capacity: Assessing the origins, evolution,*

    and impact of Race to the Top. Washington, D.C.: American Enterprise Institute.

McGuinn, P. J. (2012). Stimulating reform: Race to the top, competitive grants, and the Obama

    education agenda. *Educational Policy, 26*, 136-159.

McLaughlin, M. W. (1990). Embracing contraries: Implementing and sustaining teacher

    evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher*

    evaluation: Assessing elementary and secondary school teachers. (pp. 403-415).

    Newbury Park, CA: Sage Publications.

McLaughlin, M. W., & Pfeifer, R. S. (1988). *Teacher evaluation: Improvement, accountability,*

    and effective learning. New York, NY: Teachers College Press.

Mead, S., Rotherham, A., & Brown. R. (2012). *The hangover: Thinking about the unintended*

    consequences of the nation's teacher evaluation binge. Washington, D.C.: American

    Enterprise Institute.

Mintrop, H., & Sunderman, G.L. (2009). Predictable failure of federal sanctions-driven

    accountability for school improvement – and why we may retain it anyway. *Educational*

    Researcher, *38*(5), 353-364.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for*

    educational reform. Report of the Washington, DC: U. S. Department of Education.

National Comprehensive Center for Teacher Quality. (2012). Databases on State Teacher and

    Principal Evaluation Policies (STEP Database and SPEP Database). Retrieved from:

    http://resource.tqsource.org/stateevaldb/Compare50States.aspx

Natriello, G. (1990). Intended and unintended consequences: Purposes and effects of teacher

    evaluation. In J. Millman & L. Darling-Hammond (Eds.), *The new handbook of teacher*

evaluation: Assessing elementary and secondary school teachers. (pp. 35-45). Newbury

    Park, CA: Sage Publications.

National Education Association. (2012). New policy statement on teacher evaluation and

    accountability – adopted as amended. Retrieved from http://pareonline.net/Home.htm

    http://www.nea.org/grants/46326.htm

No Child Left Behind Act (NCLB). Pub. L. 107-110 (2002).

Osborne, J. W. (2000). Advantages of hierarchical linear modeling. *Practical Assessment,*

    Research, and Evaluation, 7(1). Retrieved from http://pareonline.net/Home.htm.

Performance Evaluation Reform Act, Senate Bill 315, Public Act 96-0861, (2010).

Peterson, K. D. (1990). Assistant and assessment for beginning teachers. In J. Millman & L.

    Darling-Hammond (Eds.), *The new handbook of teacher evaluation: Assessing*

    elementary and secondary school teachers. (pp. 104-115). Newbury Park, CA: Sage

    Publications.

Ragin, C. C. (1999). The distinctiveness of case-oriented research. *Health Services Research,*

    34(5), 1137-1151.

Sachs, J. (2001). Teacher professional identity: Competing discourses, competing outcomes.

    Journal of Education Policy, 16(2), 149-161.

Schein, E. H. (1996). *Organizational culture and leadership* (2nd ed.) San Francisco, CA: Jossey-

    Bass Publishers.

Schneider, A. L., & Ingram, H. M. (1990). Behavioral assumptions of policy tools. *Journal of*

    Politics, 52(2), 510-529.

Smith, M.S., & O'Day, J. (1991). Systemic school reform. In *Politics of education association*

    yearbook 1990 (pp. 233-267). New York, NY: Taylor and Francis, Ltd.

Smylie, M. A., & Smart, J. C. (1990). Teacher support for career enhancement initiatives: Program characteristics and effects on work. *Educational Evaluation and Policy Analysis, 12*(2), 139-155.

Spillane, J. P. (2000). A fifth-grade teacher's reconstruction of mathematics and literacy teaching: Exploring interactions among identity, learning, and subject matter. *The Elementary School Journal, 100*(4), 307-330.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387-431.

Stein, S. J. (2004). *The culture of education policy.* New York, NY: Teachers College Press.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly, 63*(3), 224-237.

Stiggins, R. J., & Duke, D. (1988). *The case for commitment to teacher growth: Research on teacher evaluation.* Albany, NY: State University of New York Press.

Students First. (2012). *Policy agenda: A challenge to states and districts: Policies that put students first.* Sacramento, CA: Author.

Superfine, B.M. (2008). *The courts and standards-based reform.* New York, NY: Oxford University Press.

Superfine, B.M. (2011). Stimulating school reform: The American Recovery and Reinvestment Act and the shifting federal roe in education, *Missouri Law Review*, *76*(1), 82-134.

Superfine, B. M., Gottlieb, J. J., & Smylie, M. A. (2012). The expanding federal role in teacher workforce policy. *Educational Policy, 26*, 58-78.

Swap, W., Leonard, D., Shields, M., & Abrams, L. (2001). Using mentoring and storytelling to transfer knowledge in the workplace. *Journal of Management Information Systems*, *18*(1), 95–114.

Teddlie, C., & Tashakkori, A. (2003). Overview of contemporary issues in mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research*, (pp. 3-50), Thousand Oaks, CA: Sage.

Tschannen-Moran, M. (2009). Fostering Teacher Professionalism in Schools: The Role of Leadership Orientation and Trust. *Educational Administration Quarterly*, *45*, 217.

U.S. Census Bureau. (2011). *Back to school 2011-2012.* Retrieved December 1, 2011 from http://www.census.gov/newsroom/releases/archives/facts_for_features_special_editions/cb11-ff15.html

U.S. Department of Education. (2009a). *Partners in reform: Remarks of Arne Duncan to the National Education Association.* Retrieved from http://www2.ed.gov/news/speeches/2009/07/07022009.html

U.S. Department of Education. (2009b). *State fiscal stabilization fund factsheet.* Retrieved from http://www2.ed.gov/programs/statestabilization/index.html

U.S. Department of Education. (2010a). *Race to the Top application for phase 2 funding.* Washington, DC: Author.

U.S. Department of Education. (2010b). *Race to the Top program guidance and frequently asked questions.* Washington, DC: Author.

U.S. Department of Labor. (1965). *The Negro family: The case for national action.* Washington, DC: U.S. Government Printing Office.

Venters, M., Hauptli, M. V., & Cohen-Vogel, L. (2012). Federal solutions to school fiscal crises: Lessons from Nixon's failed national sales tax for education. *Educational Policy, 26,* 35-57.

Vinovskis, M. A. (2009). *From a Nation at Risk to No Child Left Behind: National education goals and creation of federal education policy.* New York, NY: Teachers College Press.

Wise, A. W., & Darling-Hammond, L. (1984). Teacher evaluation and teacher professionalism. *Educational Leadership*, *42*(4), 28–35.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1985). Teacher evaluation: A study of effective practices. *The Elementary School Journal, 86*(1), 60-121.

Yin, R. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.

Teacher Perspectives on Evaluation Survey[1]

The purpose of this survey is to gain a greater understanding of the factors that influence how teachers view new teacher evaluation policies. New teacher evaluation systems are being created in response to Senate Bill 7 and the Performance Evaluation Reform Act (PERA). These laws require a new system of teacher evaluations that include the use of student growth data as part of the teacher's evaluation. Recently, changes have been made to the teacher evaluation system in your district as part of this process.

Before you fill out the answers, it's important to read the brief instructions at the beginning of each section. Your answers are absolutely confidential. **Your individual answers will only be seen by the researcher, and they will never be shared with your school leaders or district leadership. Your answers will only be reported when aggregated with the results of many other teachers.** Please read each question carefully and select the response that most closely matches your opinion.

You are eligible for this study if you are a full-time K12 teacher in your district.

1.   How familiar are you with the teacher evaluation requirements of the teacher evaluation system [TES]?

        Not at all familiar.……..…………..….………………………………..………|_|
        Slightly familiar .……..……………..….…………………………………………|_|
        Moderately familiar………..……………………………………………………...|_|
        Very familiar ……..………..….…………..……………………………………..|_|

2.   How familiar are you with the changes made to the teacher evaluation system [TES] so that it will comply with recent teacher evaluation legislation in Illinois?

        Not at all familiar.……..…………..….………………………………..………|_|
        Slightly familiar .……..……………..….…………………………………………|_|
        Moderately familiar………..……………………………………………………...|_|
        Very familiar ……..………..….…………..……………………………………..|_|

---

[1] This survey is identical to the one distributed to River District teachers, except the district identity has been concealed.

For the following four statements, please indicate the extent to which you agree with each statement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
3. In general, I am satisfied with TES.	○	○	○	○	○
4. TES can be fairly applied to all teachers.[2]	○	○	○	○	○
5. TES address almost all important aspects of teachers' work.	○	○	○	○	○
6. TES increases my satisfaction with working in my school.	○	○	○	○	○
7. TES increases my satisfaction with working in my district.	○	○	○	○	○

For the following four statements, please indicate the extent to which you agree with each statement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
8. TES supports my professional growth. [3]	○	○	○	○	○
9. TES improves the feedback I receive from my principal.[4]	○	○	○	○	○
10. The teacher evaluation requirements of TES improve the quality of instructional conversations with my principal and other teachers. [5]	○	○	○	○	○
11. TES leads to improved student performance.[6]	○	○	○	○	○
12. TES helps me to improve my teaching.	○	○	○	○	○

---

[2] Modified from Consortium on Chicago School Research. (2013). *My Voice, My School Teacher Survey*. Chicago, IL: Author.

[3] Ibid.

[4] Ibid.

[5] Ibid.

[6] Ibid.

For the following five statements, please indicate the extent to which you agree with each statement.

		Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
13.	TES is used to dismiss unsatisfactory teachers. [7]	○	○	○	○	○
14.	TES is used to help struggling teachers.[8]	○	○	○	○	○
15.	TES is used to recognize and reward excellent teachers. [9]	○	○	○	○	○
16.	TES is used to raise student achievement.	○	○	○	○	○
17.	TES is used to improve the quality of schools in my district.	○	○	○	○	○

18. Please rank the top three sources of information you have used to learn about how the recent TAP changes will affect you. (Randomize)

District Office Communications ..……………….……...…………………….|_|
News media ..……….…………….……...…………………………………….|_|
TES website ..………..………….…...……………...………………………..|_|
School leaders..………..…………….………………………..……….|_|
Teachers union………..………….…...…………………………………..|_|
Other …………………..……………….……………………………..|_|

If Other, please specify _____

---

[7] Modified from Consortium on Chicago School Research. (2013). *My Voice, My School Teacher Survey*. Chicago, IL: Author.

[8] Ibid.

[9] Ibid.

The next section will ask some questions the measures used by your district. Please indicate the extent to which you agree with each statement.

		Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
19.	State standardized test scores will accurately reflect student learning.	○	○	○	○	○
20.	State standardized test scores will accurately reflect teacher performance.	○	○	○	○	○
21.	The TES Framework is a valid and accurate way of assessing teacher quality.	○	○	○	○	○
22.	The TES Framework is addresses all aspects of teachers' work.	○	○	○	○	○

For the following set of questions, please indicate the extent to which you agree with each statement.

		Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
23.	TES recognizes the professional nature of teaching. [10]	○	○	○	○	○
24.	The ISBE teacher evaluation requirements will rely too heavily on standardized tests.[11]	○	○	○	○	○
25.	TES relies too heavily on classroom observations. [12]	○	○	○	○	○
26.	TES includes the aspects of teaching that I think are the most important.	○	○	○	○	○
27.	Overall, state standardized tests are good indicators of teacher quality.	○	○	○	○	○
28.	Teachers are the most important in-school factor in student success.	○	○	○	○	○
29.	Teachers are the most important overall factor in student success.	○	○	○	○	○

---

[10] CCSR. (2013). *My Voice, My School Teacher Survey.* Chicago, IL: Author.
[11] Ibid.
[12] Ibid.

The next section will ask you some questions about your school's response to changes to the Teacher Appraisal Program. Please remember that all of your responses will be kept confidential. Please indicate the extent to which you agree with each statement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
30. My school leader(s) clearly and carefully explained how the ISBE teacher evaluation requirements will impact TAP at my school.	○	○	○	○	○
31. My school leader(s) gave me and other teachers an opportunity to ask questions about changes to TES.	○	○	○	○	○
32. I feel like I will be involved in the implementation of the TES changes at my school.	○	○	○	○	○
33. I have had all of my questions about the ISBE teacher evaluation requirements answered.	○	○	○	○	○
34. I had enough opportunities to learn about how the ISBE teacher evaluation requirements will affect me.	○	○	○	○	○

The following questions will ask you about the person that will be, or is currently, evaluating you. Please think about this person as you answer these questions.

35. Who will be your primary evaluator?

        Principal ..…………………………………………...………….….……..…….…|_|
        Assistant principal ..…………………………………...…………….…….……|_|
        Department chair ..……………………...............……………….…...……|_|
        Central office administrators ……...…………………………….….……..…|_|
        Other ..……………………………...…………………………………….…...…….|_|

        If Other, please specify: _____

For the following questions, please think about the person that will be <u>your evaluator</u> and indicate the extent to which you agree with each statement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
36. My evaluator is a credible source of feedback. [13]	○	○	○	○	○
37. I have a cooperative working relationship with my evaluator. [14]	○	○	○	○	○
38. I trust my evaluator. [15]	○	○	○	○	○
39. My evaluator has a supportive interpersonal manner. [16]	○	○	○	○	○
40. My evaluator is knowledgeable about the technical aspects of teaching. [17]	○	○	○	○	○
41. My evaluator is knowledgeable about my particular classroom. [18]	○	○	○	○	○
42. My evaluator is fair.	○	○	○	○	○

---

[13] This question is modified from Stiggins and Duke's (1988) Teacher Evaluation Profile (TEP) questionnaire. The original question used an A-E rating scale that was presented horizontally.

[14] This question is also taken from Stiggins and Duke's questionnaire. It was modified in the same way.

[15] Also taken from Stiggins and Duke (1988).

[16] Ibid.

[17] Ibid.

[18] Ibid.

The next set of questions will ask about <u>your school</u>. Please think about the school you are currently working in when you answer these questions. Please indicate the extent to which you agree with each statement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly Agree
43. The culture of my school encourages teachers to share the struggles they have in their classrooms.	○	○	○	○	○
44. Teachers talk about instruction in the teachers' lounge, faculty meetings, or other locations in the school [19]	○	○	○	○	○
45. Teachers in this school share and discuss student work with other teachers. [20]	○	○	○	○	○
46. Experienced teachers are willing to share their practice with new teachers[21]	○	○	○	○	○
47. Teachers consider evaluation as an opportunity to get feedback on our teaching.	○	○	○	○	○

The next section of questions will ask about your motivation(s) for teaching, your goals as a teacher, and your experience as a teacher. Please select the answer that most closely describes you.

48. Please rank the top three items that motivate you as a teacher (Randomize) *1 indicates the item that most strongly motivates you to 3 is the third most motivating item*

> Addressing social inequities……………....………………………..…..………|_|
> Being recognized for my work…………. ………………………….……|_|
> Earning tenure …………...........................…………………………..….……|_|
> Having my contract renewed …………....……………………..…..….……|_|
> Helping students succeed…………………………………………..……|_|
> Achieving my career goals| …………………..……………………….……|_|
> Achieving my professional development plan …….…………...………..……|_|

---

[19] CCSR. (2013). *My Voice, My School Teacher Survey.* Chicago, IL: Author.
[20] Ibid.
[21] Ibid.

49. Please rank the top three things that you think are the purpose of education? *1 indicates the most important purpose of education to 3 indicates the third most important purpose of education*

    Developing citizens for a democratic society……………. ………………..…|_|
    Increasing social justice in society …...……………….…..…..………….....……|_|
    Helping students to develop as individuals …..,,,… …....………….…...…..……|_|
    Preparing students to the workforce…………………………......…..……….|_|
    Other.…………………….…………………….…..…………….…………….…..…….|_|

    If Other, please specify: _____

50. How much of an influence do you think a teacher has over student learning, relative to other factors such as home life, parents, etc.?

    No influence ……....……………………...…..…………………….....…|_|
    Slight influence ...………………...…………...…..………………….....…|_|
    Some influence ...………………..………….…..…………………….....…|_|
    Largest influence ...……………………...….……………………….....…|_|

51. Including this year, how many years have you been teaching in your current <u>school</u>?

    _____

52. What is the name of your current <u>school</u>?

    _____

53. What is the current grade level or content area that is your primary responsibility?

    _____

54. Including this year, how many years have you been teaching in your current <u>district</u>?

    _____

55. Including this year, how many years have you been teaching <u>in total</u>?

    _____

56. Do you have tenure?

    Yes……………………….…………………….……….……………..…….|_|
    No …………………………. …………………..………………….……|_|

57. Do you plan on remaining in education for the remainder of your career?

        Yes……………………………………………………………………..……|_|
        No (skip to #59)……………………...……………………..………………|_|
        Maybe/Unsure……………………………...…………………..………………|_|

58. Which of the following best describes your career goals in education?

        Continue teaching in my same school and content/grade………………..….……|_|
        Continue teaching but in a different school …………. ………...……………|_|
        Continue teaching but in a different district…..……….. ………...……….……|_|
        Continue teaching but in a different grade/content area…..………..….…….…|_|
        School leadership position………......................................………….…..…|_|
        District leadership position…….……...………|_|
        Education policymaking …………….........………………………..….…….…|_|

59. Thinking about where you are in your career right now, rank the following groups in the order with which you identify them. *1 = the group you most identify with to 3= the group you identify with third most* .

        Your school site……………...….…….…|_|
        Community or neighborhood your school serves …………. ………..……….|_|
        Teachers union………………………………………….………..…..……|_|
        Teach for America ......……..….……...............................…..……|_|
        Other alternative certification program......….…..……...............…..……|_|
        University-based teacher education program.........……………….…..….……|_|
        National Board Certification Program ....……………………….…..….……|_|

60. What statement best describes your route to teaching?

        University-based education program……………............................….……|_|
        Alternative certification program …………. ………..……………….……|_|
        Teach for America…....……………………………………….……|_|
        Other …...……….....................…………………………………|_|

        If Other, please specify_____

61. In the 2012-2013 school year, what ISBE rating did you receive?

        Excellent ………………………………….........................….….……|_|
        Proficient …………. ………...............................………………….…|_|
        Needs Improvement…...………...…………………………………|_|
        Unsatisfactory …...………....................................………………|_|

62. In the 2012-2013 school year, what TES rating did you receive?

Distinguished …………………………….....................…..………|_|
Proficient …………. ……..……|_|
Basic…………….…..….................……………………………|_|
Unsatisfactory …..……….....................………………………|_|

63. Of the two ratings you received in the 2012-2013 school year, which rating do you think most accurately reflects the quality of your teaching?

TES ………….. ……..............................………………….....…|_|
ISBE…..………......….......................……………….…………|_|
About the same …...………...................………………………|_|
Neither …..……….....................…………………………….…|_|

The following questions will ask you about your opinions and interest in politics. Please read each question carefully and select the response that most closely matches your opinion.

64.   Generally speaking, how would you describe your political affiliation?[22]

Republican ....………………………………………|_|
Democrat ....……………………………………….…|_|
Independent ....…………………..……………………|_|
Other ....……………………………………………|_|

If Other, please specify _____

65. In general, how would you describe your political orientation or views?[23]

Extremely liberal ....…………………………………|_|
Liberal.....................……………………………………………|_|
Slightly liberal ....……………….…………………|_|
Moderate, middle of the road ....……………………|_|
Slightly conservative...........……………….……………………|_|
Conservative …....................…………….………………………|_|
Extremely conservative ....………...………………|_|

---

[22] This question is taken from the ANES Public Affairs Profile Survey of 2010.

[23] This is also taken from the ANES Public Affairs Profile Survey of 2010.

66. In general, how interested are you in politics and public affairs?[24]

Not at all interested ....…………………....…..……………………|_|
Slightly interested ....………………….…..……………………|_|
Somewhat interested ....……....……………..……………………|_|
Very interested .....……………...…………....……………..……...…………………|_|

This final section will ask some demographic questions. Please remember that all responses will be confidential and will not be shared with anyone other than the researcher.

67. What is your year of birth?

_____

68. What is your gender?

Male…....………...……………...…………….…..………………………………….|_|
Female ...………...…..………...…………………………...…………………………|_|

69. What is your race? (*Select all that apply*)

White ………………...……………………………………………………......…|_|
Spanish/Hispanic/Latino …………………………….…..…………....……....….|_|
Black or African-American …………………...……………...……….….…….|_|
American Indian or Alaskan Native …………….....…………….……………….|_|
Asian …………………...…………………………...…………………….……….|_|
Other …………………...……………...……………...…………………….……….|_|

If Other, please specify _____

70. What is the highest degree you've received?

Associate's degree…………………...……..……………………...……………………|_|
Bachelor's degree ……………...……..………………….…..…………………….|_|
Master's degree …………..…………...………………………...…………………….|_|
Professional degree (MD, DDS, DVM, LLB, JD, DD) ………………...…..…...….|_|
Doctorate (Ph.D. or Ed.D.) …....….…….…………………...…………………….|_|

Thank you very much for your time! If you are taking this survey at a meeting, please place the survey in the envelope at the front of the room. If you are taking this survey on your own, please mail the completed survey to the researcher using the provided stamped envelope. If you are taking this survey online, no further action is needed.

---

[24] Also taken from the ANES Public Affairs Profile Survey.

If you are willing to be contacted for a confidential follow up interview with the researcher, please fill out the following information.

Name: _____

Email address: _____

 Phone number: _____

Thank you for your participation!

If you would like to be entered into a raffle to win one of three $50 Amazon gift cards, please click HERE

Interview Protocol – Selected Teachers Based on Quantitative Analysis[25]

The purpose of this interview is to talk about how you formed your opinions of the recent legislation regarding teacher evaluation in Illinois. We'll begin by talking about your overall opinion of the legislation and its requirements for teacher evaluation systems and then I'm going to ask some questions about how you developed your opinion of PERA and SB 7.

1. Tell me, in general, what do you think of the teacher evaluation system in your district?
   a. Prompt: I can see from the survey results that you felt that…

2. What do you think the teacher evaluation system is designed to accomplish?
   a. Follow up: Do you think it will accomplish that goal?

3. How do you think the teacher evaluation system impacts your teaching?
   a. Prompt if needed: Do you see any benefits to the teacher evaluation system in your district?

4. How do you think the teacher evaluation system impacts your commitment to teaching in your district?

5. How do you think the teacher evaluation system impacts your desire to remain a teacher?

6. And finally, how do you think the teacher evaluation system affects student learning?

7. Can you tell me about your experiences with evaluation? Under what other evaluation systems have you been evaluated? Can you tell me about your experiences with those? (look for most vivid examples)
   a. Prompt: Can you tell me about a particularly good experience? What made it good? What about a particularly bad one? What made it bad?

---

[25] This survey is identical to the one distributed to River District teachers, except the district identity has been concealed.

8. From the survey, I can see that the person conducting your evaluation is your principal/assistant principal/department head etc. Is that still the case?

9. How do you think your relationship with the person evaluating you affects the evaluation system?
   a. Prompt: I can see from the survey that you have a BLANK relationship with your evaluator. How does this influence how you think about your next round of evaluations?

10. Is there anyone at the district level, in the teachers' union, or outside the system who has played a role in how you view the evaluation system? If so, who is that person and how did that person affect your view of the evaluation system?

11. Why do you think there is a policy emphasis on teacher evaluation right now?

12. Why do you think most new teacher evaluation systems incorporate test scores?
   a. Follow-up: How do you think the use of test scores will affect teacher evaluations?

13. From where did you first learn about the ISBE teacher evaluation requirements? Did you think that was a reliable and accurate source? Did that source provide enough information?

14. What other sources did you use to learn about the ISBE teacher evaluation requirements? Are there particular types of information you get from each source?

15. How are the results from your evaluation used, based on your understanding the teacher evaluation system?

16. How do you think the results from the evaluations *should* be used, if at all?

17. What kind of stakes, if any, do you think should be attached to teacher evaluations?

18. If you could develop a teacher evaluation system for your district, what would you like to see that evaluation system accomplish?
    a. Follow up: What elements would you include in an evaluation system?

NAME:	Jessica Joy Gottlieb
EDUCATION:	B. S., Aerospace Engineering, University of Southern California, Los Angeles, California, 2006
	M.A., Education, Loyola Marymount University, Los Angeles, California, 2008
	Ph.D., Policy Studies in Urban Education, University of Illinois at Chicago, Chicago, Illinois, 2014
TEACHING EXPERIENCE:	Department of Educational Policy Studies, University of Illinois at Chicago: Social Foundations of Education, 2011-2013, Contemporary Controversies in U.S. Schools, 2012
HONORS:	Dean's Scholar Fellowship, University of Illinois at Chicago, Chicago, Illinois, 2013-2014
	David L. Clark Seminar Participant, University Council for Educational Administration, San Francisco, CA, 2013
PUBLICATIONS:	Superfine, B. M., & Gottlieb, J. J. (In Press). Teacher evaluation and collective bargaining: The new frontier of civil rights law. *Michigan State Law Review*.
	Superfine, B. M., Gottlieb, J. J**.,** & Smylie, M. A. (2012). The expanding federal role in teacher workforce policy. *Educational Policy*, *26*(1), 58-78.
PRESENTATIONS:	Bulkley, K. E., & Gottlieb, J.J**.** (2014, April). "Images of Teachers: Policy Design and the Social Construction of Teachers". Accepted for the American Educational Research Association Annual Meeting. Philadelphia, Pennsylvania.
	Gottlieb, J. J. (2013, April). "The Role of Politics and Science in Abbott XXI". Accepted for the American Educational Research Association Annual Meeting. San Francisco, California.
	Browne, P., & Gottlieb, J. J**.** (2012, November). "Teach for America: Charity Education in Post-Crisis Situations". Paper presented at the 13th Annual Curriculum and Pedagogy Conference. New Orleans, Louisiana.
	Gottlieb, J. J**.,** & Browne, P. (2012, October). "Theories of Action: An Examination of Teach for America." Paper presented at the Mid-

Atlantic Educational Research Conference, Philadelphia, Pennsylvania.