



PARDEE RAND GRADUATE SCHOOL

CHILDREN AND FAMILIES
EDUCATION AND THE ARTS
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INFRASTRUCTURE AND
TRANSPORTATION
INTERNATIONAL AFFAIRS
LAW AND BUSINESS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
TERRORISM AND
HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from www.rand.org as a public service of the RAND Corporation.

Skip all front matter: [Jump to Page 1](#) ▼

Support RAND

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore the [Pardee RAND Graduate School](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND electronic documents to a non-RAND website is prohibited. RAND electronic documents are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the Pardee RAND Graduate School (PRGS) dissertation series. PRGS dissertations are produced by graduate fellows of the Pardee RAND Graduate School, the world's leading producer of Ph.D.'s in policy analysis. The dissertation has been supervised, reviewed, and approved by the graduate fellow's faculty committee.

District-Union Collaboration on Teacher Evaluation Reforms

Case Studies of Three School Districts in
California

Beth Katz



District-Union Collaboration on Teacher Evaluation Reforms

Case Studies of Three School Districts in California

Beth Katz

This document was submitted as a dissertation in September 2015 in partial fulfillment of the requirements of the doctoral degree in public policy analysis at the Pardee RAND Graduate School. The faculty committee that supervised and approved the dissertation consisted of Jennifer McCombs (Chair), Rita Karam, and Julia Koppich.

Supported by the Susan Way-Smith Grant in Education.



PARDEE RAND GRADUATE SCHOOL

The Pardee RAND Graduate School dissertation series reproduces dissertations that have been approved by the student's dissertation committee.

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (<http://www.rand.org/publications/permissions.html>).

Published 2015 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665
RAND URL: <http://www.rand.org/>
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Summary

Teacher evaluations serve a dual purpose: 1) to hold teachers accountable for the quality of their instruction and student learning; and 2) to motivate and inform improvements in their practice. Given that teachers are considered to be the most important influence on student learning within the school environment, it is not surprising that policymakers look to teacher evaluations as a tool for improving student achievement. The policy landscape surrounding teacher evaluation in the U.S. is rapidly changing. Though reforms may be forwarded by state mandates, federal incentives, and funding from private foundations, implementation occurs at the district level, and in the 45 states where teachers unions are permitted, teacher evaluation procedures may be subject to collective bargaining. It follows that local stakeholders determine, to a large extent, the success of these reform efforts. If efforts to implement new teacher evaluations are to be successful – and ultimately effective – policymakers must understand when and why school district leaders, local teachers unions, and other key players collaborate, and how their interactions influence the reform process and its outcomes.

This dissertation presents three case studies of school districts in California that are working to reform their teacher evaluation policies. My research aim was to identify the conditions that led to successful development and adoption of teacher evaluation policies, as well as the factors that have contributed to challenges in those districts. I created a logical framework to outline the reform process and provide a canvas for telling the story of each case study district and for comparing cases. I explored the roles various stakeholders played over the course of district reform efforts and the evolving relationships between the districts and their teachers unions. I identified several factors that influenced the relationships between teachers unions and districts during the reform process and the policy outcomes of their efforts. Identifying the conditions that enable or impede collaboration can assist policymakers in making decisions that will encourage collaboration and minimize conflict, thus increasing the likelihood that their efforts will produce the desired outcomes. Based on my analysis, I present recommendations for policymakers, practitioners, and for further research.

Table of Contents

| | |
|-------------------------------------------------------------------|------|
| Summary | iii |
| Figures..... | xi |
| Tables..... | xiii |
| Acknowledgments..... | xv |
| Abbreviations..... | xvii |
| Chapter One: Introduction | 1 |
| Policy Problem | 1 |
| Teacher Evaluation Reform Efforts..... | 2 |
| Research Aims..... | 4 |
| Research Questions | 5 |
| Overview of This Report..... | 6 |
| Chapter Two: Background on Teacher Evaluation Policies | 8 |
| Purpose of Teacher Evaluations | 8 |
| Evaluations and Personnel Decisions..... | 9 |
| Measures of Teacher Effectiveness | 10 |
| Student Growth Measures | 11 |
| Classroom Observations..... | 14 |
| Teacher Evaluation Policy Landscape..... | 16 |
| Federal Incentives | 16 |
| State and District Policies..... | 18 |
| Bumps in the Road | 19 |
| California’s Context | 20 |
| Chapter Three: Teachers Unions and Education Reform | 22 |
| History of Teachers Unions..... | 22 |
| The National Labor Relations Act..... | 22 |
| Birth and Development of Teachers Unions | 23 |
| Industrial v. Reform Unionism..... | 24 |
| Collaborative Bargaining | 25 |
| Teachers Unions and Reform Efforts | 26 |
| Roles..... | 26 |
| Teachers Unions and Student Achievement..... | 29 |
| Theories on Labor-Management Interactions and Relationships | 31 |
| Walton and McKersie (1991)..... | 31 |
| Gates and Vesneske (2011) | 33 |
| Susan Moore Johnson (1984)..... | 34 |
| Benefits of Collaboration | 35 |
| Collaboration Case Studies | 38 |

| | |
|----------------------------------------------------------------------------|----|
| Chapter Four: Conceptual Framework..... | 40 |
| Reform Process Framework | 40 |
| Key Constructs | 44 |
| Roles in the Reform Process | 44 |
| Labor-Management Relationships | 46 |
| Influential Factors | 48 |
| Chapter Five: Research Design..... | 50 |
| Case Study Districts | 50 |
| California School District Characteristics | 51 |
| Selection Criteria..... | 52 |
| District Sample and Recruitment | 53 |
| Document Collection..... | 55 |
| Stakeholder Interviews | 55 |
| Stakeholder Recruitment and Selection..... | 55 |
| Interview Protocol | 56 |
| Analytic Approach..... | 57 |
| Step 1: Construct Case Descriptions | 58 |
| Step 2: Apply Structural Codes to the Text..... | 58 |
| Step 3: Identify Cross-Case Themes | 60 |
| Step 4: Analyze the Themes..... | 61 |
| Limitations..... | 62 |
| Chapter Six: Case Study District Overview..... | 63 |
| Existing Teacher Evaluation Policies | 63 |
| California Requirements..... | 63 |
| Case Study District Procedures | 64 |
| Common Concerns | 66 |
| District Overviews..... | 68 |
| Laurel | 69 |
| Juniper | 69 |
| Sumac | 71 |
| Organization of Results | 72 |
| Chapter Seven: Setting the Stage..... | 74 |
| Stakeholder Roles..... | 76 |
| Initiators | 78 |
| Conveners..... | 79 |
| Contributors and Supporters..... | 79 |
| Stakeholder Relationships | 80 |
| Influential Factors..... | 81 |
| Leaders Provided Momentum for the Work..... | 82 |
| New Superintendents Set the Tone for Labor-Management Interactions | 84 |
| District Culture Affected Stakeholder Willingness to Consider Change | 84 |
| A Carefully Structured and Cooperative Approach Facilitated Progress..... | 86 |

| | |
|---------------------------------------------------------|-----|
| Timing Makes a Difference..... | 86 |
| Chapter Eight: Approach to Reform | 89 |
| Stakeholder Roles..... | 91 |
| Facilitators and Communicators..... | 93 |
| Contributors..... | 94 |
| Advisors | 94 |
| Stakeholder Relationships | 95 |
| Influential Factors..... | 98 |
| The Role of the Superintendent..... | 98 |
| Joint-Leadership and Broad Representation..... | 99 |
| Topics On The Table..... | 100 |
| Time Invested In Goal-Setting | 102 |
| Planning Ahead – Timelines and Clear Expectations | 102 |
| Chapter Nine: Hammering Out the Details..... | 104 |
| Stakeholder Roles..... | 105 |
| Facilitators and Communicators..... | 106 |
| Participants | 106 |
| Advisors and Observers..... | 107 |
| Stakeholder Relationships | 107 |
| Influential Factors..... | 110 |
| Leadership Turnover Can Pose Challenges..... | 110 |
| District Culture and Contract Negotiations | 111 |
| Compromising on Reform Features | 112 |
| Committee Structure, Planning, and Persistence..... | 113 |
| Chapter Ten: Moving Forward | 115 |
| Stakeholder Roles..... | 117 |
| Facilitators and Communicators..... | 119 |
| Implementation Leads and Support..... | 119 |
| Contributors..... | 120 |
| Stakeholder Relationships | 120 |
| Sumac | 120 |
| Juniper | 123 |
| Influential Factors..... | 125 |
| Leadership and Relationships..... | 125 |
| Communication and Coordination | 127 |
| Approach: Pilot Before Permanent..... | 130 |
| Engaging Stakeholders in Quality Improvement | 131 |
| School Culture..... | 133 |
| Chapter Eleven: Conclusions and Recommendations | 135 |
| Research Question 1 | 135 |
| Superintendents | 137 |
| Assistant Superintendents of HR..... | 137 |

| | |
|------------------------------------------------------------------|-----|
| Teachers Union Leaders | 138 |
| Conclusions | 140 |
| Research Question 2 | 140 |
| Research Question 3 | 142 |
| Approach to Reform | 144 |
| Leadership | 146 |
| Reform Features | 149 |
| District Characteristics and Culture | 150 |
| Timing and External Events | 152 |
| Recommendations for Policymakers and Practitioners | 153 |
| Organizing and Managing Teacher Evaluation Committees | 154 |
| Building Relationships | 155 |
| Implementation | 155 |
| Recommendations for Further Research | 156 |
| References | 158 |
| Appendix A. Characteristics of California School Districts | 167 |
| Appendix B. Interview Recruitment Letter | 168 |
| Appendix C. Stakeholder Interview Protocol | 169 |
| Appendix D. Laurel Unified School District Case Summary | 173 |
| Stage 1: Consider Change | 173 |
| Motivation for Reform | 173 |
| Decision to Consider Change | 174 |
| Committee Recruitment | 175 |
| Stage 2: Set Goals | 176 |
| Current State of Reforms | 179 |
| Stakeholder Relationships | 180 |
| Appendix E. Juniper Unified School District Case Summary | 182 |
| Stage 1: Consider Change | 182 |
| Motivation for Reform | 182 |
| Decision to Consider Change | 182 |
| Committee Recruitment | 183 |
| Stage 2: Set Goals | 183 |
| Stage 3: Define and Discuss Options for Change | 185 |
| Stage 4: Come to Agreement | 186 |
| Setbacks | 187 |
| Stage 5: Begin Implementation | 190 |
| Stage 6: Continue Implementation | 192 |
| Stakeholder Relationships | 193 |
| Appendix F. Sumac Unified School District Case Summary | 195 |
| Stage 1: Consider Change | 195 |
| Motivation for Reform | 195 |
| Decision to Consider Change | 196 |

| | |
|-------------------------------------------------|-----|
| Committee Recruitment | 196 |
| Stage 2: Set Goals..... | 196 |
| Stage 3: Define and Discuss Options | 199 |
| Stage 4: Come to Agreement..... | 200 |
| Stage 5: Begin Implementation | 202 |
| Stage 6: Continue Implementation | 205 |
| Stakeholder Relationships | 206 |
| Appendix G. Teacher Evaluation Committees | 208 |

Figures

| | |
|----------------------------------------------------------------|-----|
| Figure 4.1 Reform Process Framework | 40 |
| Figure 4.2 Within-Stage Analysis..... | 43 |
| Figure 6.1 District Reform Timeline | 68 |
| Figure 7.1 Reform Activities – Setting the Stage | 74 |
| Figure 8.1 Reform Activities – Approach to Reform | 89 |
| Figure 9.1 Reform Activities – Hammering Out the Details | 104 |
| Figure 10.1 Reform Milestones – Moving Forward | 115 |

Tables

| | |
|---------------------------------------------------------------------------------------------|-----|
| Table 3.1 McDonnell and Pascal’s “Response Options for Teacher Organizations” | 28 |
| Table 3.2 Walton and McKersie’s “Attitudinal Components of the Relationship Patterns” | 32 |
| Table 4.1 Teacher Evaluation System Reform Process, by Reform Stage and Key Activities | 42 |
| Table 4.2 Key Constructs by Research Question..... | 44 |
| Table 4.3 Levels of Involvement | 45 |
| Table 4.4 Description of Stakeholder Roles | 45 |
| Table 4.5 Characteristics of Labor-Management Relationships..... | 47 |
| Table 5.1 Distribution of School Districts in California By Enrollment | 51 |
| Table 5.2 Distribution of School Districts in California by Teachers Employed..... | 52 |
| Table 5.3 Descriptors for Select School District Characteristics..... | 54 |
| Table 5.4 District Sample Characteristics..... | 54 |
| Table 5.5 Stakeholders Interviewed by District..... | 56 |
| Table 5.6 Structural Codes..... | 59 |
| Table 5.7 Reform Characteristics Code Descriptions..... | 60 |
| Table 6.1 District Teacher Evaluation Policies Prior to 2010 | 65 |
| Table 7.1 District Stakeholder Roles in Considering Change | 77 |
| Table 7.2 Teachers Union Stakeholders Roles in Considering Change | 78 |
| Table 8.1 Committee Activities and Established Goals..... | 90 |
| Table 8.2 Stakeholder Roles in Goal-Setting..... | 92 |
| Table 10.1 Teacher Evaluation Pilots | 116 |
| Table 10.2 Stakeholder Roles in Implementation Year 1 | 118 |
| Table 11.1 Stakeholder Roles Observed by Stage | 136 |
| Table 11.2 Individuals in Leadership Roles by District and Stage..... | 136 |
| Table 11.3 Superintendent Roles by District and Stage | 137 |
| Table 11.4 Assistant Superintendent of Human Resources Roles by District and Stage | 138 |
| Table 11.5 Union President Roles by District and Stage..... | 139 |
| Table 11.6 Union Representatives Roles by District and Stage | 139 |
| Table 11.7 Characteristics of Labor-Management Relationships Over Time | 141 |
| Table 11.8 Influential Factors by Stage | 143 |
| Table A.1 Demographic Characteristics of California School Districts, 2011-2012 | 167 |
| Table A.2 Distribution of School Districts in California by Largest Ethnic Group | 167 |
| Table G.1 District Teacher Evaluation Committees | 208 |

Acknowledgments

I first must thank the 29 participants from my three case study districts, who were generous with their time, candid in their responses, and so clearly committed to improving the quality of education for their students. Their work inspired this dissertation. Next, I am truly grateful for the support of my dissertation committee. Jennifer McCombs (chair) and Rita Karam at RAND not only offered tremendous guidance on this dissertation journey, but also provided me with opportunities to work with them on a number of research projects. Without those experiences, I would not have been prepared to tackle my own dissertation. My third committee member, Julia Koppich, provided a valuable outside perspective and much needed expertise in the field of labor-management relations in education. Without her decades of research on teachers unions and numerous publications for me to read, I may never have pursued this topic in the first place.

I am deeply indebted to Doug Smith and the Susan-Way Smith Dissertation Fund for providing monetary support for my work, and also for many engaging conversations along the way. I appreciate all the members of the Pardee RAND Graduate School faculty and staff who supported me through these often arduous years as a graduate student and who provided many cookies and other treats that lifted my spirits, including Gery Ryan, Jeffrey Wasserman, Rachel Swanger, and Susan Marquis. In particular, I would like to thank several RAND researchers with whom I have had the privilege of working with and learning from, including Charles Goldman, Jill Cannon, Brian Stecher, Abby Robyn, and Lindsay Daugherty.

I have been in school continuously for the past 28 years, as a student or a teacher, and sometimes both. There is no question that I owe a great deal of my success to my family and friends. To my incredible parents, who not only have supported me in everything I have ever done, but who have inspired thousands of their own students during their four decades as teachers. To say that their commitment to education serves as an inspiration for my work is an understatement. Lastly, to my husband, Drew, who endured this journey with me. I would have never made it through without his constant support and encouragement.

Abbreviations

| | |
|--------|----------------------------------------------------|
| AFL | American Federation of Labor |
| AFT | American Federation of Teachers |
| AP | Advanced Placement |
| API | Academic Performance Index |
| ARRA | American Recovery and Reinvestment Act |
| AS | Assistant Superintendent |
| AYP | Adequate Yearly Progress |
| BTSA | Beginning Teacher Support and Assessment |
| CBA | Collective Bargaining Agreement |
| CELDT | California English Language Development Test |
| CGTL | Center on Great Teachers and Leaders |
| CMO | Charter Management Organization |
| CORE | California Office to Reform Education |
| CSTP | California Standards of the Teaching Profession |
| CTA | California Teachers Association |
| DCPS | District of Columbia Public Schools |
| EC | Evaluation Committee |
| ED | U.S. Department of Education |
| EL/ELL | English Language Learner |
| EOC | End of Course |
| ESEA | Elementary and Secondary Education Act |
| FRPL | Free and Reduced Price Lunch |
| HR | Human Resources |
| IBB | Interest-Based Bargaining |
| LAUSD | Los Angeles Unified School District |
| LCAP | Local Control Accountability Plan |
| LCFF | Local Control Funding Formula |
| LIFO | Last-In First-Out |
| MOU | Memorandum of Understanding |
| NBPTS | National Board for Professional Teaching Standards |
| NCLB | No Child Left Behind |
| NCTQ | National Center on Teacher Quality |
| NEA | National Education Association |
| NGO | Non-Governmental Organization |
| NLRA | National Labor Relations Act |

| | |
|------|---------------------------------|
| NLRB | National Labor Relations Board |
| PAR | Peer Assistance and Review |
| PD | Professional Development |
| PLC | Professional Learning Community |
| RTT | Race to the Top |
| SGM | Student Growth Measure |
| SGP | Student Growth Percentile |
| SIG | School Improvement Grant |
| SLO | Student Learning Objective |
| SPED | Special Education |
| TIF | Teacher Incentive Fund |
| TOSA | Teacher on Special Assignment |
| TURN | Teacher Union Reform Network |
| UFT | United Federation of Teachers |
| UP | Union President |
| VAM | Value-Added Model/Measure |

Chapter One: Introduction

My dissertation explores how teacher evaluation reform efforts developed in three school districts in California, with the goal of understanding the mechanisms through which these reforms evolve. The policy landscape surrounding teacher evaluations is rapidly changing. Between 2009 and 2012, two-thirds of states amended their teacher evaluation policies (NCTQ, 2012). By 2014, 41 states and the District of Columbia had instituted a requirement that teacher evaluations be informed by some objective measure of student growth (NCTQ, 2015). States often adopted these policies in order to qualify for federal funding or satisfy the demands of their constituents. Many districts and schools that have adopted new policies are still piloting them or working to phase-in new systems over time. While the short-term outcomes of reform efforts are slowly becoming clear, the long-term outcomes of newly implemented policies – namely the effects of those policies on teacher quality and student achievement – will take considerable time to realize. In the meantime, a focus on the process through which reforms are being developed and adopted is both feasible and worthwhile.

In this dissertation, I aimed to illuminate the roles various stakeholders play over the course of these reform efforts – when they become involved and the nature of their involvement – and to identify how their individual and collective actions influence the reform process and policy outcomes. Through semi-structured interviews, I identified the internal and external influences that have acted as enablers of or barriers to successful reform efforts in those districts. Following this analysis, I highlighted strategies that encourage the engagement and cooperation of unions, districts, and other key stakeholders throughout the reform process.

Policy Problem

Improving the quality of education in the U.S. has long been a national policy priority. The publication of *A Nation At Risk* by the National Commission on Excellence in Education in 1983 has been widely credited for bringing attention to the state of public education and for mobilizing diverse stakeholder groups to address the institution's deficiencies and inefficiencies. The report found that American students were behind their international peers and that student achievement in the U.S. had actually been declining, not improving over time (Gardner, 1983). However, despite numerous attempts to reform public education in the last 30 years, the performance of American students still lags behind other developed nations. In an analysis of student performance in 2012, the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (PISA) found that 15 year-olds in the U.S. ranked 27th in math, 17th in reading, and 20th in science out of a group of 34 industrialized nations (OECD, 2014).

Domestically, there are growing concerns about educational equity. A persistent achievement gap exists between White and minority students and between students from high and low-income families. On the 2012 National Assessment of Educational Progress (NAEP), White students outscored Black and Hispanic students in reading by 21 points or more, on average. In math, White students scored 25 points or more, on average, above Black students, and the White-Hispanic gap was at least 17 points for each age group. Though the racial achievement gaps have narrowed since 1971, they persist and remain substantial (U.S. Department of Education, 2012b). Similar gaps are seen between students who do and do not qualify for the National School Lunch Program, which indicates that higher-income students are outperforming their peers from lower-income families. The OECD also found a correlation between socioeconomic status and student performance on the 2012 PISA, estimating that socioeconomic status can explain 15% of the variation in student performance in the U.S., which is similar to the average for all 34 OECD countries (OECD, 2014).

In addition to the achievement gap, there is a troubling education attainment gap in the U.S.. For the 2012-2013 school year, the public high school graduation rate for White students was 87%, while the rates for Black and Hispanic students were 71% and 75%, respectively (U.S. Department of Education, 2015c). Though the overall graduation rate for all public school students was 81%, economically disadvantaged students graduated at a rate of 73%. Similar trends are seen at the post-secondary level. Among the cohort of students entering college in 2007, the four-year college graduation rates were 43% for White students, 30% for Hispanic students, and only 21% for Black students (U.S. Department of Education, 2014).

A Nation At Risk also highlighted concerns about the nation's teaching force. Since then, additional research has shown that teachers are the most important influence on student achievement within the school environment (Sanders, Wright and Horn, 1997). Further, the cumulative effect of having less effective teachers for multiple years in a row is particularly problematic, causing students to fall behind – and stay behind – their peers taught by highly effective teachers (Sanders and Rivers, 1996). As a result, teachers have become the target of public scrutiny. Recently, conversations about student achievement have turned to discussions about teaching quality, and education stakeholders began to ask many questions, including: 1) How effective are America's teachers; 2) How central is their performance to the problems surrounding student achievement and educational equity; and 3) How could policy be leveraged to improve teacher quality?

Teacher Evaluation Reform Efforts

Education stakeholders are continuously developing, implementing, and revising education policies. Though diverse stakeholder groups – government officials, education professionals, and numerous not-for-profit organizations – share many goals, they are often at odds over specific policy solutions. Policymakers can manipulate the school environment in several ways. Common

targets of education reform efforts include changes to school structure and curriculum, as well as the implementation of accountability systems. While improvements to student learning can be approached in a number of ways, recently, many reform efforts have focused on redesigning teacher evaluations as a mechanism for improving teacher quality and raising student achievement.

Improving teacher quality is the goal of several education reforms. Though teachers have always been evaluated by their supervisors¹ – usually through classroom observations – the status quo of teacher evaluation systems in the U.S. is widely considered to be inadequate (Darling-Hammond et al., 2011). The 2009 publication of *The Widget Effect* brought attention to the state of teacher evaluation policies in the U.S., revealing that existing systems were not distinguishing effective teachers from ineffective teachers. That study of four states revealed that most districts evaluated teachers using binary ratings – satisfactory or unsatisfactory – and that in those districts, 99% of teachers were rated satisfactory (Weisberg et al., 2009). In order to improve their practice, teachers need individualized and meaningful feedback, as well as recognition for exemplary work. Telling a teacher that he or she is “satisfactory,” without additional information, is an inadequate form of feedback.

Since the late 2000s, teacher evaluation reforms have been catalyzed by financial incentives, state laws, court rulings, and pressure from non-governmental interest groups and the general public. Many stakeholders, including the Obama administration, have advocated for the use of student growth measures (SGMs) to determine, in part, assessments of teacher effectiveness, and are incentivizing states and districts to adopt those policies. The Department of Education’s Race to the Top competition awarded funds to 18 states, the District of Columbia, and 21 school districts that committed to revising their teacher evaluation systems (U.S. Department of Education, 2013). At the same time, philanthropic organizations have been providing substantial support for education reforms, with contributions totaling \$684 million from 2000 to 2008. Since then, the Bill and Melinda Gates Foundation alone committed an additional half a billion dollars to support specific education reforms (Koppich and Esch, 2012), including the use of student achievement data in calculating measures of teacher effectiveness. In 2011, the Department of Education began offering states flexibility from some requirements of the Elementary and Secondary Education Act (ESEA) that were put in place by the No Child Left Behind Act (NCLB) of 2001. In order to receive an NCLB waiver, states had to commit to several reforms, including new teacher evaluation systems that rely on SGMs as a significant factor and are used to inform personnel decisions. Currently, 42 states and D.C. have been granted ESEA flexibility (U.S Department of Education, 2015a).

¹ Teachers may also be evaluated by their peers in districts with Peer Assistance and Review (PAR) programs. In this alternative evaluation system, master teachers or peer observers evaluate and provide feedback to other teachers. They may focus on teachers identified as in need of intervention. Toledo, Ohio, was the first district to employ a PAR program in 1981. Since then, a few other districts have adopted Toledo’s model, including Cincinnati, Ohio; Poway, California; Rochester, New York; and Dade County, Florida (Goldstein, 2007).

In response to these and other incentives, two-thirds of states changed their teacher evaluation policies between 2009 and 2012 (NCTQ, 2012). As of 2014, 35 states require that objective measures of student achievement be a significant or preponderant criterion in determining teacher effectiveness, and an additional six states and D.C. specify that some objective measure of student achievement be used as a criterion. Additionally, 43 states and D.C. require that teacher evaluations use more than two rating categories (NCTQ, 2015). California is among the five states that have no requirement specifically related to measures of student growth. The state has made no recent changes to its 44-year-old teacher evaluation policy and still requires that teachers be rated as only satisfactory or unsatisfactory.

Many school districts are reforming their teacher evaluation systems in advance of or separately from state mandates, and some of these local education agencies are viewed as leaders in those reforms. Other districts are preparing to comply with state mandates – and not without difficulty. Some efforts have been successful; others have stalled at various stages in the reform process. Still others have not yet engaged in serious teacher evaluation reform efforts. In general, districts have some discretion in the type of evaluation system they develop or adopt, as long as they adhere to any basic legal requirements set by their state. In the 45 states where teachers unions are permitted, teacher evaluation policies may be subject to collective bargaining with local unions. Teachers unions have an interest in shaping teacher evaluation policies, as they have an interest in upholding the standards of the teaching profession. They may be concerned about proposals to tie evaluation results to other high-stakes decisions, such as compensation or tenure. Basing these decisions on teacher effectiveness threatens to erode contract provisions teachers unions have long protected, which may include a salary scale based on experience and qualifications. Teachers are especially concerned about tying high-stakes decisions to student test scores or other measures of student achievement, as the research on whether or not this approach is appropriate has been mixed. Unions must consider how the proposed policies will affect their members. At the same time, state and local policymakers should engage the unions representing their teachers when proposing policy changes. Teachers, who will be directly affected by those changes, may have valuable contributions to make in crafting those policies. Further, as noted above, changes to teacher evaluation policies need to be collectively bargained in many states, and thus would require the approval of both districts and unions.

Research Aims

Administrators of local public education systems are afforded limited resources to address numerous and complex problems. Given that enormous public and private resources are being directed at reforming teacher evaluations, it is in the best interest of all stakeholders that those resources be used efficiently and effectively. Though teacher evaluation reforms may be forwarded by state mandates, federal incentives, and funding from private foundations, implementation occurs at the district level. It follows that local stakeholders determine, to a large

extent, the success of these reform efforts. Not surprisingly, several school districts have encountered challenges in developing and implementing new teacher evaluation policies (Pogodzinski, Umpstead, and Witt, 2015; Aldeman and Chuong, 2014; Max et al., 2014; Donaldson and Papay, 2012; White, et al., 2012). Often these challenges seem to stem from a lack of collaboration among stakeholders. In states where teacher evaluation policies must be collectively bargained with local teachers' unions, the development and adoption of reforms may take longer than in states where those policies are not subject to collective bargaining. Even when reforms have already been adopted, implementation may stall due to resource constraints, insufficient time to prepare, or limited stakeholder support.

Prior research on teacher evaluation has primarily focused on the evaluation system itself, including the development and evaluation of specific measures of teacher effectiveness, while less work has been done to understand the process through which those systems are developed and implemented. If efforts to implement new teacher evaluations are to be successful – and ultimately effective – policymakers need to understand when and why school district leaders, local teachers unions, and other key players collaborate, and how their interactions influence the reform process and its outcomes. My dissertation addressed this gap in the literature by 1) describing the roles of teachers unions and school districts in the teacher evaluation reform process and by 2) objectively exploring how labor-management relationships affect education reforms adoption and implementation. My work builds upon existing theories regarding the roles of teachers unions and districts and evaluates how they apply in the specific context of teacher evaluation reforms.

Using three case studies, I explored the role of local teachers unions in the decision-making process. My aim was to identify the conditions that led to successful development and adoption of teacher evaluation policies, as well as the factors that have contributed to challenges in those districts. Describing how the teacher evaluation reform process has played out in local school districts – when it is successful and when it stalls – can help policymakers better plan their approaches to reforms or improve reform processes already underway. Identifying the conditions that enable or impede collaboration can assist policymakers in making decisions that will encourage collaboration and minimize conflict, thus increasing the likelihood that their efforts will produce the desired outcomes. The theory I develop regarding the process of district policy adoption, and the insights that I gather through my case studies can inform the policy decisions of stakeholders as they approach teacher evaluation and other reforms in the future.

Research Questions

In this dissertation, I used case studies of three districts to address three research questions, all related to the process through which teacher evaluation reforms are conceived, developed, adopted, and implemented:

1. What roles do teachers unions and districts play in efforts to reform teacher evaluations and to what extent does the involvement of these groups shift throughout the reform process?
2. How do teachers unions and districts interact and how do their relationships change throughout the reform process?
3. What influences the roles, interactions, and relationships between teachers unions and districts during the reform process and the policy outcomes of their efforts?

The underlying construct that spans all three questions is the reform process itself. I have defined this process as a series of six stages, which will be described in greater detail in Chapter 4. The stages in sequential order are: 1) Consider change; 2) Set goals; 3) Define and discuss options; 4) Come to agreement; 5) Begin implementation; and 6) Continue implementation. These stages served as the organizing principle for my data collection and analysis. The reform process is also shaped by a district's history and the status quo of teacher evaluations and labor-management relationships. Continued implementation will eventually be followed by observable outputs and outcomes; however, that is beyond the scope of this research.

My first research question focuses on two broad constructs – stakeholder roles and level of involvement – while my second question addresses their relationships and interactions. My understanding of these concepts was first informed by typologies used in the literature and then refined through the exploration of the data I collected. I used my data to understand the roles stakeholders play, how involved they are, the characteristics of their interactions, and the types of relationships they developed. These characteristics can and do vary across the stages of the reform process.

To address my third research question, I explored factors that influence the decision-making process and the results of the district's reform efforts as they emerged in the data. Internal and external factors may affect the reform process in various ways. Specifically, I identify the factors that shape the interactions among stakeholders, which in turn affect the policy outcomes of their efforts. These factors include district and union leadership, their approach to reform, the organizational culture of the district and union, and the broader policy environment that informs their actions. I collected data to understand what factors facilitate collaboration or incite conflict and what makes the reform process progress or stall.

Overview of This Report

In the following two sections, I review prior research on teacher evaluation reforms and teachers unions. I examine the evolution of collective bargaining in education and the current policy landscape surrounding teachers unions. Next, I present key findings from the literature on labor-management interactions and relationships. Following a review of the literature, I introduce the conceptual framework that guides my research and define the constructs related to my research questions. I then discuss my research methodology, describing my sample of

districts and data collection techniques. I also address the benefits and limitations of my approach.

My results are organized into five separate chapters that address the different stages in the teacher evaluation reform process. I integrate the stories of each of the case study districts with a discussion of each stage and include full case summaries in Appendices D, E, and F. Finally, I present my conclusions for each research question and offer recommendations for policymakers, practitioners, and for future research on labor-management collaboration around teacher evaluation reforms.

Chapter Two: Background on Teacher Evaluation Policies

Although they have received considerable attention in recent years, teacher evaluation policies are not new, and this is not the first time they have been at the center of a “reform movement.” However, the current debate diverges somewhat from past debates with its intense focus on the role of student achievement measures in teacher evaluations, and with related discussions on how teacher evaluations should inform decisions about compensation, tenure, and other personnel policies. Measures of student achievement represent one type of metric that can be used to assess teacher effectiveness. A robust teacher evaluation tool utilizes multiple measures, and generally includes data gathered from observations of classroom instruction. In this chapter, I review the various components of teacher evaluation systems and survey the evolving policy landscape. I also discuss some of the controversies surrounding recent and proposed policy changes. Lastly, I address the state of teacher evaluation policy in California, which will provide context for analyzing the reform efforts in my three case study districts.

Purpose of Teacher Evaluations

Teacher evaluation systems have several intended uses. Though evaluation procedures are not uniform across states and districts, they have similar components, uses, and users. Broadly, teacher evaluations are intended to: 1) hold teachers accountable for the quality of their instruction and student learning; and 2) motivate and inform improvements in their practice. The results of teacher evaluations are primarily used by teachers, school administrators, and district administrators.

In order for teacher evaluations to effect change, district and school administrators must hold teachers accountable for the quality of their instruction, as determined by their evaluations. In other words, they must get teachers to care about the results of their evaluations and subsequently act upon them. This also requires that teachers believe their evaluations are fair and accurate, and that they have the resources they need to make the recommended changes to their practice. How teachers should be held accountable is the subject of much debate. While some argue that the primary purpose of an evaluation should be to inform and encourage professional growth, evaluations may also be a tool to identify and dismiss less effective teachers. The extent to which administrators can achieve these objectives depends largely on teachers’ intrinsic motivation and perceptions of the evaluation system itself. Teachers want their students to achieve. Thus, they tend to be intrinsically motivated to deliver the best instruction they can to their students. However, if teachers do not believe that the performance measures to which they are held accountable accurately and reliably reflect the quality of their instruction, they will not fully accept the system. If, however, teachers value the feedback from their evaluations, they

may draw upon their internal motivation and change their practice to address weaknesses. Perceptions that an evaluation system is purely punitive, that it seeks to punish teachers for poor performance, rather than serve as a tool for professional growth, can undermine its ability to motivate improvements in teacher quality.

Professional growth is a longer-term objective of evaluation systems. Teachers need to be provided with continued support and professional development opportunities in order to improve their practice. Districts can respond to teacher evaluation results by dedicating resources to improvement efforts (e.g. professional development or instructional materials) or changing policies to address deficits that may exist in specific schools, or in the district as a whole. School administrators and coaches can work one-on-one with teachers to identify their weaknesses and create individualized action plans. They may also identify common areas for improvement across their staff and provide professional development, coaching, or other forms of support. Most importantly, evaluations provide feedback directly to teachers, who often tend to work in isolation. However, the quality of feedback provided may vary. Ideally, these efforts would result in useful and actionable feedback that results in increased teacher effectiveness over time.

Evaluations and Personnel Decisions

As mentioned above, there is a tension between using evaluations for accountability and as a professional growth tool (Mead, Rotherham and Brown, 2012). Teacher evaluations may be used by school and district administrators to inform internal decisions, such as dismissal and compensation. Though decisions related to reductions-in-force have traditionally been determined based on seniority, some policymakers have proposed that districts move away from the “last in, first out” philosophy and consider teacher effectiveness when deciding which individuals are laid-off. Some districts have experimented with pay-for-performance systems, where compensation may be increased based on evaluation scores, and some states have codified different types of performance-based compensation into law. Other less punitive consequences that are being tied to teacher evaluations include support for struggling teachers and leadership opportunities for exemplary teachers. Many school systems and their unions are designing “career ladders” that provide the most effective teachers with opportunities to take on leadership roles, such as peer coaches or demonstration teachers. This provides a pathway to promotion and potentially higher pay for teachers who do not want to become administrators.

Teacher evaluations are generally based on multiple measures of teacher effectiveness, which may include classroom observations and assessments of student learning. However, when high-stakes decisions are involved, teachers are particularly wary of the use of measures of student achievement. Even when teachers support the use of student performance measures to inform their evaluations, they are less supportive of using test scores for differentiated pay. In a study of four urban, high-poverty school districts, Wells (2011) found general support for the implementation of pay-for-performance systems, but less support for using student performance measures to determine compensation. Likewise, a nationwide survey of more than 1,100 teachers

found that, while a majority of teachers (54%) thought that measures of student progress over time are a good or excellent way to measure teacher effectiveness, they opposed the use of those measures for determining differentiated pay (Rosenberg and Silva, 2012). Only 10% of the teachers surveyed strongly favored using test scores to determine compensation.

Measures of Teacher Effectiveness

An evaluation system is only as good as the tools it uses to measure teacher effectiveness. While the focus of this dissertation is not on the merits of different evaluation techniques, it is important to understand the range of options districts may consider when designing their teacher evaluation policies. There are several measures of teacher effectiveness that may be included in a teacher evaluation system, including:

- Student achievement scores from standardized tests;
- Derived measures of student growth (value-added measures, student growth percentiles);
- Student learning objectives (SLOs);
- Alternative measures of student achievement (portfolios, performance assessments);
- Classroom observations;
- Peer ratings;
- Student perceptions (measured with surveys);
- Parent perceptions (measured with surveys);
- Professional contributions to the school;
- Participation in professional growth opportunities; and
- School-level student performance.

While it is generally agreed that multiple measures of teacher effectiveness should be used in evaluating a teacher, there is debate over the ideal system composition. Schools and districts may assign different weights to each of these measures when they compute a summative score, or they may let each measure stand alone. Several studies have cautioned against the use of student growth data as the sole measure of teacher effectiveness. Evaluations instead should be based on multiple measures of teacher performance, in order to mitigate the influence of the error and bias that characterizes any one measure (Steele, Hamilton and Stecher, 2010). Many aspects of effective teaching cannot be captured by the standardized tests administered to students, and many subjects are not currently assessed by standardized tests. The weight given to student achievement measures in a teacher's evaluation score is often a source of conflict among policymakers and especially between school districts and teachers unions. I review the use of two of the most common and heavily-weighted measurement tools – student growth measures and classroom observations – below.

Student Growth Measures

Since teachers are responsible for equipping students with the knowledge and skills required to demonstrate mastery of academic content standards, student performance on assessments of those standards is often used as a measure of the effectiveness of the instruction delivered by the teacher. Though the insight provided by test scores is limited, it is perhaps the most accessible and objective data source that is currently available – every student in the same grade or content level takes the same standardized test, usually consisting of multiple-choice questions.² In order to control for the many out-of-school factors that influence student achievement, raw scores are often transformed into derived measures of learning that attempt to quantify student growth or determine the value added by an individual teacher. There are different types of student growth measures (SGMs), which attempt to translate measures of student achievement into measures of teacher effectiveness. SGMs are calculated using value-added models (VAMs) which estimate a teacher’s contribution to a student’s academic growth in a given year. VAMs use statistical methods that control for various factors that may affect student learning, including previous test scores and several demographic characteristics, to predict how much growth a student should make in a given year. Discrepancies between actual growth and predicted growth may then be attributed to the teacher. Another derived measure of student growth is a student growth percentile (SGP). This measure is calculated by comparing students with their “academic peers” – other students who had similar performance on previous tests (Burnett, Cushing and Bivona, 2012). SGPs attempt to attribute changes to a student’s percentile rank on a standardized test – from one year to the next – to the teacher. An SGP greater than 50 suggests that the student outgrew his or her peer group – those students that began at a similar level. A teacher would then be evaluated by the median SGP for all students in the class. Because this is a norm-referenced criterion, 50% of students will always outperform their peers, while another 50% will always underperform their peers.

Student growth data can be aggregated at different levels: teacher, grade, school, or district. Teachers may not be evaluated solely on the performance of their own students, but also on school-wide performance, in an effort to hold teachers accountable for all students in their school. School-level SGMs may also be used as a substitute for teacher-level data in non-tested subjects or non-tested grades. Generally, only math and English test data has been used to measure student growth, as states are only required to administer annual standardized tests in those subjects in grades 3-8, and once grades 10-12. Some districts are developing alternative student growth measures, based on student work or teacher-created assessments, to better measure the effectiveness of teachers in non-tested subjects. Some districts rely on Student Learning Objectives (SLOs), end-of-course (EOC) assessments, or other commercially available

² Many states are in the process of adopting new standardized assessments that are aligned with the Common Core State Standards. These standards focus on high-order thinking and skills, and the related assessments being developed are less reliant on traditional multiple-choice test questions.

standardized tests. These performance measures may be used in conjunction with SGMs or VAMs, or in place of them. SLOs are student learning targets set by teachers at the beginning of the school year or semester and may be based on student achievement data from the previous year or pre-tests administered at the start of the current school year. These objectives generally must be approved by principals or district administrators, to ensure that they are sufficiently rigorous. While additional research is needed to determine if SLOs are accurate measures of teacher effectiveness, there is emerging evidence that the number of these objectives a teacher meets is positively associated with other measures of student achievement (Tyler, 2011).

Concerns

Many are concerned that SGMs are not sufficiently reliable or valid to be used as a component of teacher evaluations and especially for high-stakes decisions. These concerns include:

1. **Volatility of scores.** SGMs may be unstable across classes and years, and variable depending on the individual students assigned to the teachers (Newton et al., 2010). For example, a teacher's score may be different had he or she been assigned a different group of students at the beginning of the year. Since students are not usually randomly assigned to schools and classrooms, the reliability of the measure may be compromised (Steele, Hamilton, and Stecher, 2010).
2. **Insufficient controls in statistical models.** Although growth measures try to match students with "academic peers," who share similar demographic characteristics, there are many other factors influencing student achievement that are not modeled, including a student's home environment and parent involvement in their child's education (American Statistical Association, 2014).
3. **Questionable validity of standardized tests.** The validity of SGMs presupposes the validity of the test used to assess student achievement. However, standardized tests may be inadequate, as they cannot measure every aspect of student learning (Hamilton, Stecher and Yuan, 2008; May et al., 2009). Some argue that paper-and-pencil tests taken by students, which provide the data to calculate SGMs, are not "authentic" or direct measures of teacher performance (Goe, 2010). This weakens their validity as an indicator of instructional quality.
4. **Sensitivity to model specification.** Different models for calculating SGMs can produce different results (American Statistical Association, 2014). Additionally, the same tests are not available for all subjects. As a result, teachers in the same school may be assigned SGMs calculated using different types of assessments. A VAM score for one teacher may vary when different student achievement tests are used (Lockwood et al., 2007), or when different statistical methods are applied (Briggs and Domingue, 2011; Newton, 2010).

The American Statistical Association (ASA) recently cautioned against the use of VAMs for high-stakes decisions, noting that most of the variation in student test scores is due to factors outside of a teacher's control. The ASA emphasized that while VAMs measure correlation, they do not attribute causation. Further, since VAMs are calculated using only student test scores, they cannot assess other important student outcomes (American Statistical Association, 2014). Complicating the use of SGMs is the fact that nearly all states are currently in the process of transitioning to the new Common Core academic standards, which require new standardized tests (Mongeau, Felton and Butrymowicz, 2015). When the tests administered to students change, it becomes difficult to make longitudinal comparisons.

In addition to the methodological concerns, there are some practical issues with using SGMs in teacher evaluations. Practical concerns include:

1. **Difficult to accurately and clearly communicate.** SGMs may be difficult for the end-users, who are not statisticians, to interpret. It is important that the statistical error associated with a given measure be clearly communicated and taken into consideration when interpreting results (Ballou and Springer, 2015).
2. **Cost and technical requirements.** The use of SGMs requires substantial expertise and technical capacity at the district level, which can be costly and difficult to sustain (American Statistical Association, 2014).
3. **Unintended consequences.** There is some concern that the use of SGMs in teacher evaluations may discourage teachers from working in high-needs schools, since they may be penalized for poor student performance (Johnson, 2015).
4. **A narrowing of the curriculum to teach to the test.** As tests carry high stakes for teachers, many stakeholders are concerned with an increased focus on teaching to the test. There is evidence that a narrow focus on preparing students for a test results in less attention to non-tested subjects and skills (Darling-Hammond and Adamson, 2014). There is a growing standardized test "opt-out" movement among parents who are frustrated with the increasing focus on testing.
5. **Availability of data when testing requirements change.** There are also emerging concerns about the sustainability of these measures. It is unclear if states will continue to sustain the current level of commitment to standardized tests if Congress relaxes federal testing requirements, when they replace NCLB with new legislation in 2015. If states reduce the number of standardized tests they administer, then districts may need to find alternative assessments of student achievement, in order to calculate SGMs.

The practical concerns described above, coupled with the associated methodological issues, have important consequences for stakeholder support. Teachers trust these measures less than other performance measures and principals may be hesitant to use them for personnel decisions (Goldring et al., 2015; Jiang, Sporte and Luppescu, 2015).

Classroom Observations

Classroom instruction is the mechanism through which teachers influence student behaviors (Grossman et al., 2010). Observations of classroom instruction are the only tool available that allows administrators to directly evaluate a teacher's instructional practice, and thus are an essential component of an evaluation system. They are also important because a large number of teachers cannot be rated using test scores (Whitehurst, Chingos and Lindquist, 2014), and also because using multiple measures of teacher effectiveness results in more robust evaluations (Kane and Staiger, 2012; Goe, Bell and Little, 2008). Further, observations are needed to assess aspects of teacher performance that cannot be inferred from student achievement tests.

Traditionally, teacher evaluations have consisted of classroom observations conducted by school administrators, often using a simple checklist. While probationary teachers may be evaluated every year, permanent or tenured teachers may only require evaluations every two to five years and therefore, may rarely be observed by an administrator. Current reform efforts are attempting to make observations more systematic, objective, and rigorous, through the use of standards-based rubrics to aid in collecting data from classroom observations. These rubrics are divided into several categories or domains of teaching, such as planning, instruction, and classroom environment. Each domain is broken down into specific characteristics that the observers will rate according to a pre-determined scale, based on the evidence they collect in the classroom. An aggregate measure from the rubric is used in the teacher evaluation score.

An evaluation system needs to be grounded in an understanding of what constitutes effective teaching. Classroom observation tools are designed around a set of professional teaching standards. The National Board for Professional Teaching Standards (NBPTS) are commonly used as the basis for observation tools, as is Charlotte Danielson's Framework for Teaching. California has its own standards, the California Standards for the Teaching Profession (CSTPs), which some districts in the state use in their teacher evaluations. Research supports using standards-based evaluation tools to measure student learning and also support professional growth (Darling-Hammond et al., 2012). Further, classroom observations present an opportunity for administrators to provide timely, formative feedback to their teachers, which complements the summative assessment at the end of the year. Observations may be supplemented with teacher-created lesson plans and artifacts from the classroom that provide evidence of student learning. They generally involve teacher-principal conversations both before and after the formal observation. Classroom observations can provide teachers with specific and targeted feedback that cannot be extracted from student test scores alone. Teachers can then use this data to make changes in their practice.

Concerns

There are several concerns surrounding the use of classroom observations to evaluate teachers, related to their validity, reliability, and cost. These include the following:

1. **Validity of a single observation.** Some studies have found that standards-based observation scores correlate with VAM scores, and because of that, they can be considered valid measures of teacher effectiveness (Kane and Staiger, 2012; Sartain, Stoelinga and Brown, 2011). Still, while a rubric may be grounded in teaching standards supported by research, some critics are concerned that observations may not be an entirely valid measure of a teacher's effectiveness in the classroom. Since formal observations are generally scheduled, teachers may put on a "dog and pony show," and tailor their instruction to meet criteria enumerated in the rubric, rather than following their usual day-to-day practices. To address this issue, evaluation systems may require more than one observation in a given year and often mandate a certain number of unannounced or informal observations.
2. **Reliability across evaluators.** Consistency across all evaluators using the same tool is a concern. To ensure inter-rater reliability, evaluators must attend trainings and often get formally certified in using the rubric (Graham, Milanowski and Miller, 2012). However, the amount of training districts provide is highly variable. Some schools rely on principals as the sole evaluators. Other school systems use additional administrators, peer teachers, and outside evaluators to conduct the observations. Though some consider outside evaluators to be the most objective, others cite concerns about their lack of contextual knowledge about the specific school's environment, as well as their potential lack of subject matter expertise.
3. **Inherent bias and subjectivity.** No matter what observation rubric is adopted, there will always be a degree of subjectivity involved. Some research has found that teachers with higher-performing students tend to receive higher scores on classroom observations than teachers with lower-performing students (Whitehurst, Chingos and Lindquist, 2014; Chaplin et al., 2014).
4. **Required resources.** Though teachers are in their classrooms every day, making it theoretically feasible to conduct observations, the process can be extremely time-consuming. Districts must think carefully about the frequency and duration of each required observation and the number of individuals that will be conducting the observations, and make sure that school administrators have the resources available to meet those requirements, while fulfilling their other job responsibilities (Malen et al., 2011). Districts may need to hire additional school administrators or evaluators, which can be costly. Regardless of who conducts the observations, districts need to continuously train new and existing observers to maintain reliability, incurring significant life cycle costs (Graham, Milanowski and Miller, 2012). Unless adequate resources are provided, it may be difficult to sustain the human capital needed over time.

Despite these and other concerns, it is likely that the practice of conducting classroom observations will be sustained, since most stakeholders agree that they are valuable and

important. However, as discussed, they are only one component of a robust evaluation tool, which requires multiple measures of teacher effectiveness.

Teacher Evaluation Policy Landscape

Current efforts to reform teacher evaluation policies are part of a larger movement toward standards-based accountability systems, which are motivated by the belief that all groups of students should be held to and helped to achieve high academic standards. As discussed, policies that aim to hold teachers accountable for student performance have recently become the focus of efforts to improve public education in the U.S. Federal and state policy action is increasingly requiring districts to adopt new teacher evaluation systems. The status quo of teacher evaluation is often criticized because formal classroom observations – those used to determine evaluation scores – may be infrequent or because the majority of teachers receive satisfactory ratings. Further, traditional evaluation systems have failed to differentiate teachers, rating them on a binary scale of “satisfactory” or “unsatisfactory,” with no distinction for those with exemplary performance. Current reform efforts have attempted to address these deficiencies; however, the vast majority of teachers are still receiving satisfactory ratings. While accountability systems tend to focus on identifying, and sometimes dismissing, the lowest-performing teachers, teacher evaluations also serve as tools for professional growth for all teachers. In this section, I review some of the federal incentives motivating states and districts to pursue teacher evaluation reform, survey the state of teacher evaluation policies in states and districts, and discuss some of the challenges they have faced in implementing new policies.

Federal Incentives

Several federal initiatives have encouraged states, districts, and schools to establish accountability systems. The federal No Child Left Behind Act (NCLB) of 2001, which reauthorized the Elementary and Secondary Education Act (ESEA), required states to develop standards-based assessments and administer them annually to students in grades three through eight, with the goal of having all students reach proficiency on those assessments by the 2013-14 school year; a goal that has proven difficult for states to reach. As of 2003, the law required states to adopt uniform accountability systems and hold schools accountable for making adequate yearly progress (AYP) toward student achievement goals for all groups of students, including disadvantaged students and those with special needs. Schools that received Title 1 funds (which are given to schools with large populations of disadvantaged students) would be labeled as “in need of improvement,” and could even be closed for failing to reach their AYP targets. NCLB has been associated with an increasing focus on standardized testing nationwide. Additionally, NCLB required states to employ “highly qualified teachers,” as determined by credentials and certification exams. Though the definition of “highly qualified” varied by state and was not defined by measures of teacher effectiveness in the classroom, the highly qualified teacher

requirement brought attention to general concerns about teacher quality. With the resulting increase in standardized testing, states now had more data available to measure student progress, and potentially teacher effectiveness.

In 2006, the U.S. Department of Education (ED) introduced the Teacher Incentive Fund (TIF), a competitive grant program that encouraged districts to implement performance-based compensation systems in high-need schools. The grant's goal was to attract high quality teachers to teach in low-performing schools serving predominantly low-income and minority students and to reward those teachers for gains in student achievement (U.S Department of Education, 2015b). An evaluation of the 2010 TIF grantees by Mathematica noted several implementation challenges, with fewer than half of those districts implementing all of the required program components (Max et al., 2014).

In 2009, the Obama administration created the Race to the Top (RTT) competition, which offered financial incentives to states that commit to a number of reforms, including the development of standards-based assessments for students and the use of those assessments in teacher evaluations. In 2012, ED launched a Race to the Top District competition. Eighteen states, the District of Columbia, and 21 individual school districts have received Race to the Top funds (U.S. Department of Education, 2013). The Obama administration also increased funding for the existing School Improvement Grant (SIG) program and modified the eligibility requirements. SIG funding can be awarded to the lowest-performing schools in each state, through a competitive application process. Schools receiving funds must commit to adopting one of four intervention models. One option, the "transformation model," mandates that schools implement rigorous teacher and principal evaluation systems that incorporate SGMs (U.S. Department of Education, 2012).

Another catalyst of teacher evaluation reforms is the ED's initiative to offer states flexibility with respect to the accountability systems required by the ESEA, under NCLB. These waivers exempt recipients from the requirement that all students meet the 2013-2014 proficiency targets, as well as the consequences associated with not meeting those targets. In order to receive flexibility, states must institute several reforms, including new teacher evaluation systems that rely on SGMs as a significant factor and are used to inform personnel decisions. Teacher evaluation systems must also utilize at least three ratings levels, rather than a binary system. ED began granting waivers in 2011, with the first waivers expiring in 2014. Forty-two states and the District of Columbia are approved for ESEA flexibility and 24 have been granted renewals, extending flexibility beyond the original expiration date (U.S Department of Education, 2015a). Washington, which was granted flexibility in 2012, had its waiver revoked in 2014 after failing to pass legislation requiring the use of student achievement measures in teacher and principal evaluations. Both California and Iowa had their original requests denied. However, a collaborative of eight school districts in the state, who are members of the California Office to Reform Education (CORE), was awarded its own waiver in August, 2013. This was the first time waivers were granted to individual districts. Seven of the districts applied for and were given a

flexibility extension, with the warning that they were at high risk of having the waiver revoked. Since then, another district withdrew from the joint-application.

State and District Policies

According to the National Center on Teacher Quality, between 2009 and 2012, two-thirds of states changed their teacher evaluation policies (NCTQ, 2012). By 2013, 45 states had updated their legislation or rules regarding teacher evaluations; California is among the five states that have not made changes to their teacher evaluation policies. Forty-three states passed legislation that specifically addressed the use of SGMs in teacher evaluations and the majority of states planned to roll out their teacher evaluations statewide by the 2014-2015 school year (CGTL, 2013). By 2014, 35 states required that objective measures of student achievement be a significant or preponderant criterion in determining teacher effectiveness scores. Six additional states and the District of Columbia require that some objective measure of student achievement be used as a criterion. Additionally, 43 states and D.C. require that teacher evaluations use more than two rating categories (NCTQ, 2015).

Changes to other human capital policies have accompanied teacher evaluation reforms in several states. Tenure or continuing contract laws vary by state and provide protections to new teachers after they successfully complete a probationary period. In 2009, no states explicitly required districts to take student performance into account when making decisions on granting tenure. By 2013, 20 states had adopted policies mandating that student performance be considered in such decisions (NCTQ, 2014a). In addition, 29 states have articulated that teacher ineffectiveness can be grounds for dismissal (ibid).

A handful of districts acted to implement new teacher evaluation systems in advance of state legislative action. One of the first districts to pursue aggressive reforms was the District of Columbia Public Schools (DCPS) under the leadership of Chancellor Michelle Rhee. Shortly after taking office in 2007, Rhee quickly implemented a number of reforms and unilaterally decided to close several schools and fire many principals. In 2009, DCPS adopted a new teacher evaluation and compensation system, IMPACT. Under the IMPACT system, teachers are evaluated with a combination of VAM scores and classroom observations, and are assigned ratings on a four-category scale of effectiveness. IMPACT scores were tied to compensation, with the highest rated teachers earning substantial bonuses. However, the lowest-rated teachers were subject to dismissal. Between 2009 and 2012, almost 400 teachers were dismissed for poor performance (Brown, 2012). Rhee's aggressive approach to reform provoked significant backlash and stirred national controversy. She resigned in 2010. Critics were dissatisfied with the lack of stakeholder involvement in decision-making. Some have praised the reforms for its high retention rate for top-rated teachers and dismissal of poor performers, while others have criticized the reforms for lowering the overall teacher retention rate and the negative effects on teacher morale (Zubrzycki, 2012). DCPS, and other early adopters of similar teacher human capital reforms, such as Denver, Colorado and Pittsburgh, Pennsylvania continued to draw

national attention to the teacher evaluation debate. In 2009, the Bill and Melinda Gates Foundation launched the Intensive Partnerships for Effective Teaching initiative to promote the spread of teacher evaluation and related reforms, including performance-based compensation. The foundation provided six years of funding to support the development and implementation of these reforms in four Intensive Partnership Sites (IPS): Hillsborough, Florida; Memphis, Tennessee; Pittsburgh, Pennsylvania; and a group of charter management organizations (CMOs) in California.

Bumps in the Road

States and districts have faced many challenges as they have moved to adopt and implement new teacher evaluation policies. Some have faced legal challenges. Others have been mired in controversy surrounding the public disclosure of teacher evaluation scores or cheating on standardized tests. These unanticipated challenges and controversies can complicate reform efforts and delay implementation. NCTQ analyzed contracts in 120 large school districts, including the largest school district in each of the 50 states, and found that about 72% now use measures of student achievement in their teacher evaluations (Joseph, 2015). However, other research has found that implementation of newly adopted state and district policies has been slow. Despite state mandates, many districts are still not utilizing SGMs in calculating teacher ratings or using teacher effectiveness scores to make personnel decisions (Aldeman and Chuong, 2014). Because of simultaneous policy changes, including the current shift to new academic standards, states may not be making as many modifications to their teacher evaluation systems as they had originally planned. In fact, some states have decided to delay implementation of policies that were already adopted until they complete the transition to the new Common Core standards and related student assessments. According to a survey by *The Hechinger Report*, most states using new Common Core-aligned student assessments are not associating them with high-stakes decisions for students or teachers. Only thirteen states will use test scores to inform teacher evaluations in 2014-2015 and another 21 have a plan for doing so in the future (Mongeau, Felton and Butrymowicz, 2015).

Recently, the implementation of many reforms has also been initiated or thwarted by court rulings. On June 13, 2012, a Los Angeles County Superior Court Judge determined that the Los Angeles Unified School District (LAUSD) – the second largest school district in the country – was in violation of California’s 44-year-old Stull Act for not basing teacher and principal evaluations on student test scores. However, this interpretation of the law is disputed. *In Doe v. Deasy*, the judge noted that the current evaluation system is inadequate; in 2009-10, 99.3% of LAUSD received the highest rating, “yet state standardized test scores that year showed that only 45% of students performed at grade level in reading and 56% in math” (Watanabe, 2012).

As new teacher evaluation policies were being adopted, public disclosure of teacher evaluation data became the subject of vigorous debate. In 2010, and again in 2011, the *Los Angeles Times* published individual VAM scores for thousands of LAUSD elementary school

teachers, which were calculated by an independent statistician hired by the newspaper, not the school district. While some stakeholders contend that parents have a right to know how effective their child's teacher is, others maintain that the public disclosure of individual teachers' evaluation results is a violation of their privacy and may have negative consequences. Teacher advocates argue that a singular number, derived from a statistical methodology most people do not understand, should not define a teacher in the public's eye. Alone, a VAM score does not paint a complete picture of teacher effectiveness, which is best assessed with multiple measures, and can be misleading. Briggs and Dominigue (2011) recalculated the same LAUSD teachers' VAM scores using different statistical models than the analysis in the *Los Angeles Times*. Teachers' effectiveness ratings were very sensitive to the models used. Under Briggs and Dominigue's model, only 46% of teachers kept the same effectiveness rating for reading outcomes, as did 60% of teachers, when looking at math outcomes. When accounting for sampling errors, which were not reported by the *Los Angeles Times*, the authors concluded that around half of the teachers' ratings could not be distinguished from an average effectiveness rating with statistical confidence.

Since 2010, states have been considering laws that clarify what data can be made public. In 2012, Massachusetts, New Jersey, New York, Tennessee, and Utah strengthened rules protecting teacher personnel data from full public disclosure (Sawchuk and Wickner, 2013). In New York, parents are given their students' teachers' overall effectiveness ratings, though not scores on individual components of the evaluation. In Arkansas and Indiana, schools have to report average data so that individual teachers cannot be identified. As of 2015-16, Michigan will have to notify parents if their children have had "ineffective" teachers for two years in a row (Brown and Balingit, 2015). In 2014, Florida was forced to release teacher evaluation scores after the *Florida Times-Union* newspaper sued to gain access (Jester, 2014). Virginia is currently fighting a lawsuit that challenges a 2013 state law that made teacher evaluation scores confidential. Virginia has been collecting teacher SGP scores since 2011 but has found the measures to be flawed and may be looking for replacements (Brown and Balingit, 2015).

California's Context

While the majority of states have recently amended their teacher evaluation policies, teacher evaluations in California are still governed by the Stull Act (California Education Code §44662) originally passed in 1971. While the law requires that student progress toward academic standards be considered, it does not mandate that SGMs, or other specific measures of student achievement be part of teacher evaluation systems. The Stull Act provides a framework for evaluation; however, the direct components and processes of teacher evaluations systems are determined at the district level through collective bargaining with the local teachers' unions. The 2012 ruling against the LAUSD in *Doe v. Deasy*, found that the district was in violation of the Stull Act for not incorporating student achievement into teacher evaluations. While the lawsuit

put a focus on the use of standardized test scores, districts had long been considered in compliance with the law by using classroom observations. In fact, most districts in the state do not utilize measures of student achievement when evaluating teachers. One survey of districts in 2011 found that student outcomes were not being used at all in seven of the 10 largest school districts in the state (Aldeman and Chuong, 2014). The Stull Act allows districts to use any standards enumerated in the National Board Professional Teaching Standards (NBPTS) or the California Standards for the Teaching Profession (CSTPs), which were adopted in 1997 and most recently revised in 2009, as part of their teacher evaluations.

California legislators have not been receptive to proposed changes to teacher personnel policies. The Center of Great Teachers and Leaders (CGTL) found that California is the only state that does not have plans for a new teacher evaluation system, based on multiple measures, to be implemented, while all other states have planned for implementation by 2019 (Munguia, 2015). The state was denied Race to the Top funding grant and ESEA flexibility because it would not commit to linking teacher evaluations to student test scores. In the past few years, some state legislators have proposed bills that would change teacher evaluations and other personnel policies, but they have not been successful. For example, in 2013, the California Senate Education Committee rejected SB 441, which would have mandated four categories of teacher effectiveness to be used in teacher evaluations and would have increased the frequency of evaluations for veteran teachers from every five years to every three years (Byer, 2013). In 2013, California's Governor Brown vetoed AB 375, which would have streamlined the process for dismissing teachers, which is currently a lengthy process (The Times Editorial Board, 2013). However, the bill did not propose any related changes to the state's evaluation system.

Despite inaction in the California legislature, advocacy groups are pursuing change through the courts. In 2012, the advocacy group, Students Matter, brought a lawsuit against the state on behalf of nine students. *Vergara v. California* challenged five specific state statutes related to tenure, dismissal, and seniority, which plaintiffs argued made it too difficult to fire grossly ineffective teachers. The plaintiffs argued that low-income and minority students were being disproportionately taught by grossly ineffective teachers and that they were deprived of their right to "good teaching," which should be protected under the equal protection clause. In June, 2014, a Los Angeles County Superior Court judge ruled in favor of the plaintiffs, finding that the statutes in question were indeed unconstitutional (Blume and Ceasar, 2014). The state is appealing the ruling, in what will likely be a lengthy appeal process. While this lawsuit did not directly target teacher evaluation policies, its argument relies on the availability of accurate measures of teacher effectiveness. Further, a reliable teacher evaluation system will be necessary should existing personnel policies need to be replaced by laws that require dismissal decisions to be based on teacher effectiveness, rather than seniority. Should the ruling be upheld, California will need to pass legislation linking teacher tenure decisions more directly to teacher evaluations. Until then, we are likely to see more movement on teacher evaluations among California districts than in Sacramento.

Chapter Three: Teachers Unions and Education Reform

My three research questions address the roles played by teachers unions and school districts, as well as their interactions and relationships, in the context of teacher evaluation reform efforts. My conceptual framework, which will be introduced in Chapter 4, is informed by the literature on teachers unions and labor-management interactions. In this chapter, I briefly review the history of teachers unions and their roles in prior education reforms. I discuss different paradigms of unionism and comment on research that attempts to link teachers unions to student achievement. Lastly, I explore the literature on labor-management interactions and district-union collaboration.

History of Teachers Unions

The National Labor Relations Act

In response to poor working conditions and the abuse of labor in the early 1900s, as well as the resulting labor unrest, Congress passed The National Labor Relations Act (NLRA) in 1935. While public employees are not protected under NLRA, the law became the model for state legislation that later protected the rights of public sector workers, including teachers. The NLRA addressed practices that compromised the welfare of workers, such as the unwillingness of employers to recognize their workers' right to organize, as well as employers' refusal to engage in collective bargaining. The NLRA was preceded by the Norris-La Guardia Act, passed in 1932, which prohibited employers from forcing employees to sign contracts that vowed they would not join a union. This enabled unions to increase their membership and power. The NLRA went further and enumerated the rights of private sector employees³ to form labor organizations, bargain collectively, and elect a single organization to represent them at the bargaining table for the purpose of negotiating a legally binding contract that would cover wages, hours, and terms and conditions of employment. Other activities related to collective bargaining and protecting the interests of union members were also addressed. The NLRA defined unfair labor practices as actions by employers that interfere with the rights of employees. These include discouraging membership in labor organizations, refusal to bargain collectively, firing a member for filing a grievance, and forcing employees to enter into a contract. The NLRA created the National Labor Relations Board (NLRB), which was tasked with overseeing and addressing labor issues and grievances, and adjudicating unfair labor practices. The NLRB evaluates workers' petitions to strike, investigates claims of unfair labor practices, and may petition the United States district

³ The NLRA applies to private sector employers engaged in inter-state commerce. Agricultural employers are excluded.

courts to take action on violations. The NLRA still governs labor-management practices in the private sector today.

Birth and Development of Teachers Unions

Teachers associations existed long before the passage of the NLRA in 1935. The National Education Association (NEA) was formed in 1857, followed by the American Federation of Teachers (AFT) in 1916. Today, those organizations represent more than 3.2 million and 1.5 million education professionals, respectively. Since their inception, both organizations have been influential in shaping education policy and collective bargaining legislation. The NEA, which was originally the National Teachers Association, initially represented education associations in 15 states. The organization went on to lobby for the creation of a federal department of education. The NEA pushed for increased teacher pay and other benefits, such as pensions (NEA, 2014). The AFT, which is associated with the AFL-CIO advocated for teacher tenure provisions in the 1930s and was influential in the establishment of tenure rights in 17 states by the end of the decade. However, teachers associations had no formal collective bargaining rights in any state until the late 1950s (American Federation of Teachers, 2015).

Unions organized to pass legislation governing the collective bargaining rights of public sector workers in individual states. Most states modeled their own legislation governing collective bargaining rights on the NLRA and thus legislation across the states tends to be similar. Wisconsin was the first state to pass collective bargaining legislation in 1959 and other states followed in the 1960s. In response, teachers unions grew. State collective bargaining laws may include provisions that vary for different types of public workers, including teachers, police officers, and firefighters. State laws also determine the scope of collective bargaining and identify which issues may be the subject of negotiations.⁴ Thirty states and the District of Columbia require collective bargaining of teachers, 15 states allow, but do not require them to organize, and five states explicitly prohibit bargaining (NCTQ, 2014b). Twenty-four are considered “right to work” states. In those states, labor organizations are prohibited from collecting agency fees that are automatically deducted from the paychecks of all employees represented by the union, including individuals who have not chosen to register as union members. These fees cover the costs of collective bargaining and other services provided by the union. These “right to work” laws tend to curtail union membership. In thirty-five states, it is illegal for teachers to strike (Sanes and Schmitt, 2014).

In the 1960s, teachers organizations began to assert a greater role in school decision-making. At the time, education systems were controlled by local school boards and administrators, and teachers had little clout when it came to policy making. Teachers wanted input on education

⁴ As of 2014, North Carolina, South Carolina, and Virginia prohibit collective bargaining across the board for all public sector workers. While Georgia prohibits collective bargaining for public sector employees, it makes an exception for firefighters. Similarly, Texas makes exceptions for firefighters and police officers. Tennessee makes an exception for teachers (Sanes and Schmitt, 2014).

policies including pay, class size, and curriculum. In 1968, the United Federation of Teachers (UFT) led a series of teacher strikes over the course of several months in New York City and brought national attention to the changing role of teachers unions. Teachers associations transitioned from being solely professional organizations to labor unions with collective bargaining rights (Moe, 2011). One effect of unionization is that teachers assumed a greater role in policymaking (Cooper and Sureau, 2008). In her case studies of schools in six districts, Johnson (1984) found that in general, the introduction of collective bargaining enhanced the rights of teachers and constrained the authority of administrators, but that the effects were not that radical. It promoted the centralization and standardization of policies, and the standardization of personnel policies instituted more defined hierarchies and grievance procedures. Initially, contract negotiations were limited in scope and focused on wages, hours, and terms and conditions of employment. Their objectives included reducing favoritism and abuse from administrators, establishing reasonable work expectations, and creating fair evaluation procedures (Johnson and Kardos, 2000). Teachers unions advocated for provisions that protected their own interests, as well as the interests of students, including increased preparation time and smaller class sizes. Importantly, teachers unions wanted to ensure the equitable treatment of members and protect teachers from arbitrary decisions by administrators. Contracts addressed that concern by making district policies consistent; every teacher was bound to the same rules.

Industrial v. Reform Unionism

Teachers unions began as professional associations and later adopted collective bargaining strategies that were modeled off of their industrial counterparts. State collective bargaining laws resembled the NLRA, which was written to address the needs of labor in private sector industries. The resulting form of collective bargaining is often referred to as “industrial” or “traditional” unionism. It is characterized by adversarial labor-management relationships, where unions make demands and managers resist, and a focus on equity and uniformity for those employees covered under a contract (Johnson and Kardos, 2000). While negotiations are generally adversarial, both sides must compromise in order to agree to the terms of a contract. Many of the industrial-style tactics that teachers unions adopted from industry are still used today and some argue that this approach has been an impediment to education reform. In fact, the NLRA explicitly discourages labor-management cooperation, in order to prevent unions from being co-opted by management, which would ultimately limit their ability to advocate for the rights of workers (Kerchner and Koppich, 2000).

Many teachers unions have been transitioning away from traditional industrial-style unionism toward a more professional unionism, which emphasizes collaborative decision-making over adversarial relationships (Kerchner and Koppich, 1993). Unions are increasingly focusing on teacher quality and upholding professional standards, rather than maintaining a narrow focus on upholding due process. Unions that chose to move toward a “new unionism” or “reform

unionism,” recognized that the public good, not just the interest of teachers are at stake when considering reforms. With this approach, labor and management acknowledge their shared responsibility for high quality classroom instruction and work together to solve problems and reach long-term goals (Kerchner and Koppich, 1993). Even when both parties commit to collaboration, interactions can still be adversarial at times. Periods of conflict are to be expected, even when collaboration is a priority.

This era of reforms began in the 1980s. Some teachers unions started to explore alternatives to industrial-style bargaining. In 1985, Albert Shanker, the president of the AFT, gave an influential speech promoting reform unionism, in response to growing concerns about the state of the education in the U.S., such as those raised in *A Nation at Risk*. A decade later, in 1997, Bob Chase, president of the NEA, made a similar call for a “new unionism.” He encouraged union locals to work with school administrators on improving quality, rather than focusing exclusively on working conditions and job security. The adoption of peer review programs is an example of a major change from the traditional spread of “bread-and-butter” contract issues (Johnson and Kardos, 2000). In 1996, teachers unions from 12 school districts came together to form the Teacher Union Reform Network (TURN) and restructure the role of teachers unions. The organization promotes labor-management collaboration and urges its 200 affiliated union locals to focus on teacher professionalism and be involved in shaping teaching and learning policies.

Kerchner and Koppich (2000) recognized that there are many barriers to reform unionism. Unions may lack the organizational capacity to take on new roles or implement reforms. Also problematic are the existing cultures of teachers and school administrators, in which standards-based reforms, such as teacher evaluations, are often perceived as mechanisms for removing teachers. Lastly, there are generally hostile policies that are not conducive to this reform culture, which could be addressed by amending labor laws (Kerchner and Koppich, 2000b). Still, progress has been made in many districts that have adopted more progressive forms of bargaining. Recognizing that there is further work to be done, they call for more substantive policy research on the role of unions and policy options for improving education quality.

Collaborative Bargaining

There are two approaches to bargaining described in the literature. Traditional, industrial-style negotiations fall into the category of “distributive” bargaining, whereas more progressive forms of negotiations are described as “integrative” (Walton and McKersie, 1991). Distributive bargaining has its roots in conflict, with two sides fighting over limited resources. Integrative bargaining, on the other hand, is an approach to problem-solving grounded in common interests. Fisher and Ury proposed the method of principled negotiation, which is also referred to as the “win-win” approach, as an alternative to traditional positional bargaining, which follows the distributive model. While this method applies to negotiations in any setting, it has been associated with progressive forms of unionism. Rather than bargaining over positions, parties

focus on areas of shared interest and look for opportunities for mutual gain. The method has four principles (Fisher, Ury and Patton, 1991):

- “Separate the people from the problem;”
- “Focus on interests, not positions;”
- “Invent options for mutual gains;” and
- “Insist on using objective criteria.”

Koppich and Kerchner (2000) describe the purpose of collaborative bargaining as follows:

The goal of collaborative bargaining is to dampen labor-management conflict by constructing a system in which each side recognizes the interests of the other as well as their mutual interests.

Many progressive school districts have adopted collaborative approaches to negotiations, which require different types of labor-management relationships. Two of my case study districts (Laurel and Sumac) adopted an integrative approach to negotiations called “interest-based bargaining” (IBB), which was first described by Fisher and Ury. Unlike traditional negotiations, the two sides come to the bargaining table with a list of interests, rather than a pre-determined set of positions, such as a specific salary increase. The two districts included in this study remarked that they prefer IBB to adversarial negotiations.

Teachers Unions and Reform Efforts

Though teachers unions are important players in any education reform, there has been relatively little attention given to their roles, especially at the district level, and to their interactions with superintendents and school boards. Additionally, discussions about the role of teachers unions have often been one-sided – either overly critical or very supportive (Kerchner and Koppich, 2000). Further, research focusing on the role of teacher unions in the context of collective bargaining has overlooked the work they do outside of contract negotiations (Fowles and Cowen, 2015). The purpose of this dissertation is not debate whether or not teachers unions should exist, but to better understand the roles that unions may place in specific reform efforts at the district level, not limited to their place at the bargaining table. This section reviews prior research on teachers unions’ roles in education reform, followed by a discussion of research linking union activities to student outcomes.

Roles

Since teachers gained the right to organize in the 1960s, teachers unions, as organizations, have had two main roles. They must 1) negotiate for benefits for their members, effectively acting as political interest groups, and 2) meet member demands to maintain membership (McDonnell and Pascal, 1988). Teachers unions wield influence through collective bargaining and political activities, and are considered to be one of the most powerful interest groups in state and national politics. Unions have a clear stake in education policy reform. Cowen and Strunk

(2015) review three perspectives on the role of unions in education policy. The first two perspectives are described by those who are generally critical of teachers unions. One is that unions are “rent-seekers,” focused of increasing compensation and improving working conditions for their members (Hoxby, 1996). The second perspective is that teachers unions are political interest groups that, in addition to seeking improved working conditions, will oppose any policy that weakens their organizations (Moe, 2011). A major source of a union's power is its political influence. Union-endorsed candidates will be more aligned with their policy preferences. As such, unions tend take an active role in local school board elections. The third perspective aligns with the paradigm of reform unionism. From this perspective, unions act as professional organizations that aim to increase teacher professionalism by supporting teachers and improving their working conditions (Mausethagen and Granlund, 2012; Kerchner and Koppich, 2000a; Kerchner, Koppich and Weeres, 1997; Bascia, 1994; Kerchner and Koppich, 1993). Unions that adopt practices aligned with reform unionism still must advocate for benefits for their members, but are also willing to move beyond traditional bread-and-butter contract issues and assert a great role in policy-making.

The roles unions play in reform efforts and the manner in which they react to different types of proposals may vary. McDonnell and Pascal (1988) identified three common ways that teachers unions respond to reform:

- Oppose/resist policies.
- Adapt/accommodate new circumstances imposed on them.
- Shape policy by playing an active role.

Unions must assess the risks, costs, and payoffs of each strategy, both internally and externally. There is an internal risk in advocating for positions not aligned to the sentiment of the union membership. Should union leaders not please their membership, there is the risk that members will elect a new organization to represent them, terminate affiliation with any union, or vote current leadership out of office. However, progressive leaders may intentionally push their members to adopt new positions. Externally, unions must be cognizant of their public image. For example, if they lobby for benefits that have little public support, they may gain a negative public image. This is risky, because if they are not respected by the general public, union leaders may have limited access to the policymakers who vote on and fund reforms. Thus, union leaders must balance internal and external support to be effective at shaping education policy. McDonnell and Pascal summarize the risks and payoffs associated with the three response options in the table below.

According to McDonnell and Pascal, and as shown in Table 3.1, union leaders who defend the status quo may face less internal risk, but a greater risk of losing public support. Accommodating reform proposals advanced by others has the highest payoff probability, but the future benefits of the proposed policies are necessarily unclear. With this approach, the union must convince its membership that some degree of change is necessary, while still protecting

members from the provisions of the proposed reforms that would be most unfavorable. The most risky approach is for unions to assume leadership roles, as it can be difficult for union leaders to advocate for reforms that may not benefit all teachers. Union leaders must successfully argue that preserving the status quo is a worse option than a new approach, recognizing that some members will be more amenable to change than others. McDonnell and Pascal noted that, while the probability of success is uncertain, especially in the short-term, there is great potential for long-term gains if reforms shaped by unions are implemented successfully.

Table 3.1 McDonnell and Pascal's "Response Options for Teacher Organizations"

| Option | Risk | | Payoff | |
|-------------------------------------|-----------------|--------------------|------------------------|--------------------------------------------------------|
| | Internal | External | Probability of Payoff | Level of Payoff |
| Defend status quo | <i>Low</i> | <i>High</i> | <i>Low</i> | <i>Low</i> |
| Accommodate reform proposals | <i>Moderate</i> | <i>Moderate</i> | <i>High</i> | <i>Uncertain to moderate</i> |
| Shape new approaches | <i>High</i> | <i>Low to high</i> | <i>Moderate to low</i> | <i>Low to uncertain (short-term); high (long-term)</i> |

Source: McDonnell and Pascal (1988), *Teacher Unions and Educational Reform*, p. 16.

McDonnell and Pascal also observed different responses to reform efforts from national, state, and local unions. At the state level, NEA and AFT affiliated unions sometimes took different positions from one another, and there was also variation in union responses across state unions with the same national affiliation. Union reactions to reform efforts may also reflect the organizational culture and nature of collective bargaining in individual districts. It can be difficult for state and local unions to be leaders in reform because of the fundamental challenge of having two roles: maintaining membership support while being an effective interest group in the external environment. They found that local unions focused on obtaining material benefits and that their membership would react negatively to union efforts to forward reforms related to "teacher professionalism," such as performance-based pay. Securing "bread and butter" contract issues was a precondition for unions accepting contract provisions favorable to teacher professionalism. However, the reform unions that have since emerged have departed from traditional unions, asserting leadership roles in reform efforts. In fact, initiating change is central to their missions (Johnson and Kardos, 2000; Kerchner and Koppich, 1993)

In an analysis of two case studies of a provincial and a state-level teachers union in Canada and the U.S., respectively, Poole (2001) observed two types of union responses to reform efforts. Poole found that unions responded reactively to externally-imposed initiatives, often wanting to preserve the status quo. However, at other times they took anticipatory approaches. These instances suggest that there may be opportunities to change the way in which unions are involved in reform. However, Poole noted that it would not be sufficient for unions alone to change their behaviors; all stakeholders must be willing to form partnerships with teachers unions (Poole, 2001).

In a previous case study, Poole (1999) observed that the teachers union had a preference for influencing policy in the early stages of its development, not after it had already been drafted, but that they may not have the professional relationships in place that would allow them to participate early on. She suggests that bypassing unions in policymaking efforts, or excluding them in the beginning stages of the process may force unions to take reactive and political roles; by doing so, unions are then seen as obstructionists. In other words, the perceived risks of including them are not as worrisome as the risks of excluding them. However, additional research is needed to understand the implications of teachers unions' involvement in education reform. Others have recognized the importance of engaging stakeholders early in the reform process, noting that such an approach is crucial for establishing relationships and building support (Minnici, 2014; Goe, Holdheide and Miller, 2011).

Teachers Unions and Student Achievement

As Susan Moore Johnson (1984) observed, "Collective bargaining, being a process rather than a product, can have both good and bad effects." While individual teachers are the most important in-school influence on student achievement, the link between teachers unions and student outcomes is unclear. Whether or not the provision of collective bargaining and even the existence of teachers unions altogether impacts educational outcomes is the topic of debate. The literature remains divided on whether or not it is possible to establish a link between the presence of unions or collective bargaining laws and those outcomes, including student achievement (Cowen and Strunk, 2015; Superfine and Gottlieb, 2014; Burroughs, 2008). Previous studies have employed different, and possibly flawed, methodologies and have, not surprisingly, arrived at different conclusions (Cowen and Strunk, 2015; Goldhaber, 2006). Some studies have used binary indicators of unionization, such as the presence of a union or the existence of a CBA, as their explanatory variables, while others have derived more nuanced measures of contract strength or restrictiveness to use in their models. Collective bargaining laws vary by state and actual union practices can vary considerably across districts and states, which can complicate this analysis (Fowles and Cowen, 2015).

In a recent comprehensive review of the literature, Cowen and Strunk (2015) found that while the evidence is mixed, there may be slightly negative effects, though further research is needed. Some studies authored by union critics have reported negative effects of collective bargaining on student achievement (Moe, 2009; Hoxby, 1996). Where negative associations were found, effects varied by district characteristics, with stronger negative impacts found in large districts and high-minority schools (Strunk and McEachin, 2011; Moe, 2009). Contract restrictiveness has also been associated with lower student achievement, as measured by district performance on California's Academic Performance Index (API), and graduation rates (Strunk, 2011; Strunk and McEachin, 2011). While causality cannot be proven, these studies assume that restrictive contracts constrain administrators, limiting their ability to make quality improvements. However, there is research refuting that assertion. Recently, Vachon and Ma (2015) found a

positive association between teachers unions and student achievement, a conclusion supported by previous studies (Eberts and Stone, 1987; Nelson and Rosen, 1996; Carini, Powell and Steelman, 2000).

Some critics maintain that teachers unions support the status quo and are not interested in meaningful reform (Winkler, Scull and Zeehandelaar, 2012). Strunk and Grissom (2010) found that stronger, more politically active unions have less flexible CBAs, which constrains the policy options available to districts, thus limiting reform efforts, especially with respect to personnel and evaluation policies. Applying this logic, other studies have attempted to link collective bargaining and educational inequities. The argument behind this approach is that collective bargaining enables teachers unions protect the seniority rights that control transfers and school placement, and that those policies burden underprivileged students with the lowest-quality teachers (Hill, 2006). This same argument, that seniority and tenure policies are responsible for perpetuating educational inequities, is at the center of recent lawsuits, such as *Vergara v California* (2014). However, other research has shown that collective bargaining does not, in fact, exacerbate teacher quality gaps (Cohen-Vogel, Feng and Osborne-Lampkin, 2013; Koski and Horng, 2007). Further, unions do not always fight against change, and they have many organizational strengths and contributions to make to teaching policy (Johnson et al., 2009; Bascia, 2005). Johnson, et al. (2009) rejected the claim that unions should be eliminated to achieve reform and actually found that union presidents are an important source of leadership. In many cases they have supported or spearheaded reforms, and have used collective bargaining as a means for improving student achievement (Vachon and Ma, 2015).

Despite the lack of agreement in the literature, there have been recent efforts by policymakers to scale back the scope of collective bargaining for teachers and weaken their unions. These efforts may be linked to education reform proposals that also address teacher evaluation policies or be part of broader attacks on all public sector unions. At least twelve states modified their collective bargaining laws between 2010 and 2014 (Superfine and Gottlieb, 2014). Some states narrowed the scope of acceptable topics for collective bargaining, while others curtailed collective bargaining rights more broadly. These states include Idaho, Indiana, Michigan, Ohio, and Tennessee, and Wisconsin. The efforts of Wisconsin Governor Scott Walker to limit teachers' collective bargaining rights helped spark an effort to recall him from office, which he survived. Lawmakers' efforts to restrict collective bargaining in Idaho and Ohio were later repealed by voters. These legislative attacks on unions may have consequences for relationships between school districts and their teachers unions, and could possibly make it more difficult for them to work together on reforms. In the following section, I review theories governing labor-management interactions and relationships, and discuss the potential benefits associated with district and union collaboration on reform.

Theories on Labor-Management Interactions and Relationships

In *A Union of Professionals: Labor Relations and Educational Reform*, Kerchner and Koppich (1993) described labor (teachers unions) and management (superintendents and school boards) as an essential partnership, because the two entities are interdependent. One party cannot change local education policy without cooperation from the other, when changes must be collectively bargained. Though each party may be naturally self-interested, they are not destined to be adversaries. The relationships between school districts and unions may be entirely collaborative, completely in conflict, or fall somewhere between those two extremes. The following section reviews various typologies of labor-management interactions and relationships from the literature on unions, some related to teachers unions, and some focused on unions more broadly. I use these models to inform my conceptual framework, which will be discussed in Chapter 4. I also review literature that explores the benefits of and precursors to labor-management collaboration and look at case studies of districts that were successful in fostering collaborative and productive relationships.

Walton and McKersie (1991)

In *A Behavioral Theory of Labor Negotiations*, Walton and McKersie (1991) took a social-psychological perspective on labor negotiations and identified four activities that are central to the negotiations process: distributive bargaining, integrative bargaining, attitudinal structuring, and intraorganizational bargaining. They used these activities as a framework for understanding the behavior of the parties involved in negotiations. Each activity has related tactics. Distributive bargaining is based on conflict. The parties compete to influence the division of limited resources. A strong focus on this approach often leads to one side getting more than the other and fewer joint gains. Integrative bargaining is a problem-solving approach where parties identify areas of common interest and opportunities for joint gain. There is no fundamental conflict between the interests of the two parties because they share common concerns. However, if one side focuses too heavily on this approach, they risk being taken advantage of by the other side. Attitudinal structuring refers to the way in which activities influence attitudes and how attitudes shape relationships. Negotiators can choose to change their attitudes in order to shape the negotiations process. Lastly, intraorganizational bargaining involves the behaviors of a negotiator within an organization and specifically to the task of reaching consensus among the organization's members. A negotiator must balance the task of coming to agreement with the other party while maintaining support from his own party.

The attitudes of each party influence the character of the labor-management relationship. Walton and McKersie (1991) proposed five labor-management relationship patterns: conflict, containment-aggression, accommodation, cooperation, and collusion. These patterns are related to the attitudes held by each party, which they define across four dimensions: 1) motivational orientation and action tendencies toward each other (competitive-individualistic-cooperative); 2)

beliefs about the other's legitimacy; 3) feelings of trust toward the other; and 4) feelings of friendliness-hostility toward the other. The authors noted that these attitudes and relationships play an important role in labor negotiations and have implications for the results of bargaining efforts (Walton and McKersie, 1991).

Table 3.2 Walton and McKersie's "Attitudinal Components of the Relationship Patterns"

| Attitudinal Dimensions | Pattern of Relationship | | | | |
|--------------------------------------------------------------------|---------------------------------------------|--------------------------|-------------------------------------|----------------------------------------------|-------------------------------------------|
| | Conflict | Containment-aggression | Accommodation | Cooperation | Collusion |
| Motivational orientation and action tendencies toward other | Competitive tendencies to destroy or weaken | | Individualistic policy of hands off | Cooperative tendencies to assist or preserve | |
| Beliefs about legitimacy of other | Denial of legitimacy | Grudging acknowledgement | Acceptance of <i>status quo</i> | Complete legitimacy | Not applicable |
| Level of trust in conducting affairs | Extreme distrust | Distrust | Limited trust | Extended trust | Trust based on mutual blackmail potential |
| Degree of friendliness | Hate | Antagonism | Neutralism-Courteousness | Friendliness | Intimacy – "Sweetheart relationship" |

Source: Walton and McKersie (1991), *A Behavioral Theory of Labor Negotiations: An Analysis of a Social Interaction System*, p.189.

The motivational orientation of either party influences labor-management interactions. A competitive orientation describes the desire to defeat the other side, even though there may be associated costs and sacrifices. This attitude is central to the conflict and containment-aggression relationship patterns. An individualistic tendency lends itself to a "hands-off" approach. Self-interests are pursued while no effort is made to either support or defeat the other party. This is characteristic of the accommodation relationship pattern. Lastly, a cooperative motivational orientation is associated with a concern for both parties and is characteristic of the cooperation and collusion relationship patterns.

Parties may have preferences for different relationship patterns and may even prefer to be in conflict. At the most negative end of the spectrum described by Walton and McKersie is the conflict relationship pattern. The two parties feel distrust and hate for each other, resulting in irrational or extreme behavior. The parties do not accept each other as legitimate, and management only interacts with the union to the extent required by law. The containment-aggression relationship is also competitive, but less negative. The parties "grudgingly" accept that the other side is legitimate, yet they are still antagonistic and distrustful and want to weaken the other side. To that end, management may try constrain the union's influence and even solicit allegiance from union members. Bargaining has a narrow focus on bread and butter issues and is characterized by power plays. This relationship resembles those related to traditional, industrial style unionism.

Between conflict and cooperation lies the accommodation relationship pattern. The motivational orientation of each party is individualistic. They accept each other's legitimacy, but

still have limited trust. They maintain some level of respect and do not try to interfere with the other side. They are neither antagonistic, nor friendly. The cooperative relationship is founded in mutual respect. The parties are both trusting and friendly and work toward the common goal of strengthening the organization; they do not want the other side to fail. Negotiations go beyond the required bread and butter issues to include other shared concerns. This relationship type is consistent with the paradigm of reform unionism in education. The collusion relationship, while also rooted in cooperative tendencies, is more intimate than the cooperative relationship. Labor and management work closely together outside of negotiations. They may form a coalition to forward the organization at the expense of a third party. The relationship is built on trading favors and an understanding that both sides are capable of blackmailing each other.

Walton and McKersie explored the various determinants of the different relationship patterns. They identified several contextual factors that can influence the labor-management relationship, including the economy, industry, politics, and community. They also explore how personality characteristics affect the level of trust and degree of friendliness between parties. A leader with an authoritarian and hostile personality may inhibit trust and cultivate a competitive relationship. Similarly, parties are less likely to be cooperative if they are suspicious of one another or have low self-esteem. Ideology may also play a role. A leader's beliefs with respect to the interests of either party can influence labor-management interactions. Walton and McKersie noted that, in general, labor-management relationships move from conflict to accommodation over time. However, the tenure of the leadership may inhibit change.

Gates and Vesneske (2011)

Gates and Vesneske (2011) evaluated labor-management interactions in the context of education. While they focused on interactions in the context of contract negotiations, the behaviors they observed capture the range of behaviors that might be expected away from the bargaining table, as well. They observed contract negotiations in three school districts experiencing conflict and interviewed representatives from both the district and teachers union. They identified five "action frames" to describe the range of interactions between district and union leadership observed during the contract negotiations: craftsmanship, stewardship, competition, solidarity, and diplomacy. These frames describe how leaders of unions and districts interacted with their constituents and what their underlying ideologies were. The authors discussed the implications of each approach in the context of four tasks related to negotiations: 1) planning; 2) managing; 3) communicating prior to bargaining; and 4) communicating during bargaining.

Stewardship is characterized by top-down communication and authority-based relationships, with a focus on efficiency and frugality. While efforts to be frugal in contract negotiations can be beneficial for a school district's bottom line, a strong emphasis on cutting costs can generate criticism for leaders. Gates and Vesneske also observed downfalls related to the prioritization of

efficiency. Attempts to move quickly through the negotiations process actually led to setbacks that slowed down the process.

Leaders who exemplified the craftsmanship approach were focused on process – upholding rules and following standard procedures. These leaders had substantial knowledge and skills related to negotiations and were able structure negotiations according to recognized procedures and norms. However, leaders who focused too much on procedure often made the process cumbersome. Gates and Vesneske note that not all individuals who sit at the bargaining table will have as much training and experience as the leaders and many will not be familiar formal procedures. As a result, they may perceive a leader who fits within the craftsmanship frame to be rigid and prone to conflict.

The solidarity approach may be pursued in response to a lack of agreement. Gates and Vesneske observed union leaders actively soliciting support from and mobilizing their membership. Organizing a strike may be consistent with this approach. Leaders use membership support as a source of power for persuading the other side to change positions. Though this can be an effective way to overcome a roadblock in negotiations, it can also result in a lack of trust and damage already fragile labor-management relationships.

The competition frame is centered on conflict and characterized by adversarial relationships. Gates and Vesneske observed district leaders taking this approach after becoming “fed up” with the union. Leaders focus on winning and maintaining their integrity while taking defensive and offensive positions. This aligns with Walton and McKersie’s (1991) description of the competitive motivational orientation, which they associate with the conflict and containment-aggression relationship patterns. In this frame, either side may see themselves as forwarding a moral imperative. However, this rigid approach may cause leaders to overlook potential opportunities for collaboration and agreement.

In contrast, the diplomacy approach puts a priority on problem-solving. Leaders work to achieve compromise while maintaining relationships. Gates and Vesneske note that these behaviors are consistent with the interest-based bargaining approach to negotiations. Both sides must be willing to work together and make an effort to understand each other. This requires communication and quick resolution of misunderstandings. However, by taking this approach, union leaders risk being seen too close to district leaders by their membership. This is a common and central challenge for union leaders that is identified across the literature.

Susan Moore Johnson (1984)

Susan Moore Johnson (1984) conducted case studies of six school districts and evaluated labor-management relationships at the district and school level. She found that the most important factor that influences labor relations is people, and the resulting interpersonal relationships. The personalities and actions of four key players largely determined the character of local teachers unions: union leaders, district administrators, principals, and union site representatives. Labor relations varied across three categories: collaborative, cooperative, and

contentious. She observed that if any one key player was combative, then the relationship was more formal and adversarial.

The “labor-management relationship” at the district level is not the same as the teacher-principal relationship observed at the school level. The teachers Johnson observed wanted to have “professional” relationships with their administrators. Most preferred to resolve issues with principals rather than initiating formal grievance procedures. However, union site representatives who chose to be more oppositional than cooperative behaviors instituted more formal labor relations. The attitude of principals provoked different union responses. If principals viewed the union as a threat, intentionally ignored the contract, or rejected the legitimacy of site representatives, then the union would take a confrontational stance. However, when principals respected and “willingly honored” the contract, then staff were more cooperative and supportive. She writes, “Collective bargaining is a people-centered process, just as schools are people-centered places” (Johnson, 1984). The relationships between teachers and principals at individual schools influence the way the contract is implemented. As a result, contract provisions and implementation varied across districts and labor practices varied at school sites, even within the same districts.

Johnson identified three other factors that influence labor relations at the district level: district size, the labor tradition of the community, and whether enrollment was growing or declining. In smaller districts, stakeholders were better able to trust each other, come to mutual understanding, and make informal agreements. In cities that had a strong tradition of support for labor, the districts had stronger unions and collective bargaining contracts; history mattered. Lastly, negotiations were more difficult in districts with declining enrollment. The threat of decreased funding resulted in a strong union focus on job protections and stringent enforcement of contract provisions. On the other hand, in districts with growing enrollment and thriving local economies, negotiations took a less belligerent tone. Their unions were open to a more flexible enforcement of provisions that could be tied to job security. Factors that did not explain variation across districts include urbanicity, years of collective bargaining, and number of previous strikes. While she found that local actors had influence over the labor relations at their schools, the effects of demographic factors were not uniform across schools.

The works of Johnson, Walton and McKersie, and Gates and Veneske all show that labor-management relationships vary across a range of characteristics and are not static. The authors identified many of the same factors that influence those relationships, with individuals being the most important. The attitudes and behaviors of key players in the district and union shape labor-management interactions. In my analysis of the case study districts, I explored the roles of key stakeholders and observed many of the behaviors identified in these previous studies.

Benefits of Collaboration

Collaborative labor-management relationships, a central tenet of reform unionism, are considered to be enablers of education reform efforts. These efforts are not always born at the

bargaining table, and successful districts and teachers unions have found ways to work together on changing policy outside of contract negotiations. One way to do this is to form committees that operate outside of negotiations and give teachers the opportunity to shape district policy-making (Behrstock-Sherratt et al., 2013; White et al., 2012; Johnson and Kardos, 2000; Koppich and Kerchner, 2000; Kerchner and Koppich, 1993; McDonnell and Pascal, 1988). As the ED has encouraged and incentivized states and districts to engage in teacher evaluation and other challenging education reforms, they have identified labor-management collaboration as a crucial prerequisite for successful reform efforts. The ED is motivated by a hypothesis that:

Collaboration is a more effective and efficient way to develop great teachers and strong instructional systems, and that it is a more sustainable approach over time than the ups and downs of adversarial relationships (U.S Department of Education, 2012).

However, it recognizes the need for further research to evaluate whether or not that hypothesis holds true. Specifically, the ED calls for research that examines different types of collaborative decision-making and the policies those processes produce. The ED also emphasizes the importance of collaboration in some of its key programs. For example, states applying for RTT funds had to first secure the support of their respective teachers unions. States that were unable to do this were disqualified from the competition.

Policymakers may be inclined to circumvent teachers unions when implementing reforms. As discussed, a few states do not allow collective bargaining and several have passed legislation to weaken public sector unions in recent years. While some have argued that this is the only viable option for moving forward, that power held by teachers unions must be drastically reduced, others have found that collective bargaining agreements themselves are not always obstacles to reform. Strunk (2009) looked at collective bargaining agreements across California and found both variation and opportunities for collaboration. Public sector unions, including teachers unions, can either facilitate or obstruct collaboration (Kearney and Hays, 1994). As Poole (1999) noted, excluding teachers unions in decision making, especially at the early stages of reform efforts, can lead to reactive or obstructionist behavior. Thus, it would behoove policymakers to take the risk of involving unions in decision making, as discounting them may result in no progress at all. Others have come to similar conclusions, observing that unions may be forced to take reactionary positions, potentially opposing reforms altogether, if they are not involved in policy design (Donaldson and Papay, 2012).

In Michigan, state law was recently changed to prohibit collective bargaining over teacher evaluations. At the same time, the state mandated the adoption of new evaluation systems, giving districts some flexibility in their design. Although districts in Michigan could have adopted new policies without consulting their teachers unions, one study found that the teachers union did have some influence in 80% of the districts interviewed (Pogodzinski, Umpstead and Witt, 2015). Additionally, administrators reported benefits of including their teachers unions in the policy design and implementation process and were satisfied with the outcomes. The authors

found a small, but positive correlation between a harmonious labor-management climate (as reported by district administrators) and the level of union involvement in that reform process. This highlights an interesting disconnect between the state lawmakers, who saw unions as an impediment to reform and thus took action to weaken their collective bargaining rights, and local administrators, who saw value in working with their teachers unions on evaluation policies, even if it was no longer in the context of collective bargaining.

Regardless of whether or not collective bargaining is permitted by state law, a hostile labor-management relationship can have negative consequences that influence the behaviors of district and union stakeholders (Koppich, 2005; Johnson and Kardos, 2000). Kerchner and Koppich (1993) identified mutual respect as a prerequisite for moving toward reform unionism and more productive labor-management agreements. Likewise, in a study of the 2008 Bellvue teachers strike in Washington, Jacoby and Nitta (2012) observed how the union strayed from the less combative strategies that are typically associated with reform unionism, and resorted to more traditional, combative tactics, in order to get a seat at the decision-making table. Jacoby (2011) also found that bypassing teachers unions did not lead to better outcomes. In fact, ignoring teachers' interests can be harmful to reform efforts, because teachers are the primary deliverers of educational services. He argued that teachers unions spend too much time responding to threats to their existence, which has taken time away from forwarding a progressive vision of education that moves beyond the antiquated "industrial model" of schools. Teachers should be involved in both research on improving the quality of education provided and direct contract negotiations.

A cooperative labor relations climate has been associated with improved organizational performance outside of the education sector, as well. In a study of 305 branches of a unionized banking system in Australia, Deery and Iverson (2004) found that cooperative labor relations can lead to a greater organizational commitment and union loyalty, resulting in increased productivity, quality of service, and decreased absenteeism. They found that productivity and service quality was positively related to perceptions that management was fair and facilitated cooperation (Deery and Iverson, 2004). With a survey of 350 large, unionized manufacturing facilities in the U.S., Cooke (1989) found that a greater intensity of joint labor-management activities, which was assessed by observing program structure, union strength, and other organizational factors, can lead to greater productivity and quality improvement.

In a 2011 survey conducted by Rosenberg and Silva (2012), teachers who described the relationship between their union and district as collaborative and trusting were more likely to express positive feelings about their most recent evaluations; in districts where the district-union relationship was characterized by conflict, teachers were 15 points less likely to agree that their recent evaluations were "done carefully and taken seriously by the administration." However, it is important to note that a trusting relationship does not guarantee progress on reform efforts. In fact, a pre-occupation with building collaborative and trusting relationships may serve as an impediment. Leaders must be committed to change. They must be willing to discuss topics that

may have been previously avoided and be prepared to navigate inevitable disagreements, rather than avoiding conflict altogether (Koppich and Kerchner, 2000). Further, there are many determinants of whether or not district reform efforts will be successful, and ultimately written into new CBAs. These may include perceived mutual interest, the relationships formed in implementing past contracts, the union's membership's preferences, and the internal relationship between union leaders and their members (McDonnell and Pascal, 1988).

Collaboration Case Studies

There are various ways that districts can work collaboratively with their teachers unions, and effective strategies are often identified through district case studies. There are many districts often cited as models of successful labor-management partnerships.⁵ Rubinstein and McCarthy (2011) conducted case studies of six districts that had fostered and sustained long-term labor-management partnerships and a district culture of collaboration. They argued that those partnerships are necessary prerequisites for lasting reform. Generally, these districts were motivated to form partnerships by a specific event, where stakeholders realized that the adversarial status quo was not sustainable or productive. These districts formed joint labor-management planning bodies that involved teachers unions in district-level decision-making. Many negotiated the structure of these committees into their CBAs or drafted MOUs, which enabled future collaboration. The partnerships were founded on a commitment to evaluating teacher quality and student achievement and finding solutions to problems. The authors found that these partnerships were enabled by the long tenure of district and union leaders. That stability helped institutionalize the culture of collaboration. In a follow-up study, the same authors suggested that labor-management partnerships led to more school collaboration and improved student performance, and that districts could benefit from increased state and federal support for those partnerships (Rubinstein and McCarthy, 2014).

Using case studies of 12 school districts conducted as part of the U.S. Department of Education's Labor Management Conference in 2011, Eckert (2011) identified many of the same themes. The case study districts were engaged in a number of policy initiatives, including teacher evaluation, compensation, and career development, which they approached collaboratively. In fact, many districts crafted joint-statements that expressed their commitment to collaboration. Underlying their work was a shared commitment to make student achievement a priority. Many of the districts formed joint decision-making or oversight bodies and wrote them into their CBAs. Their work required them to go beyond the traditional topics of collective bargaining, and some adopted more progressive collective bargaining strategies, such as IBB.

⁵ These include, but are not limited to: ABC Unified School District, California; Baltimore Public Schools; and Douglas County Public Schools, Colorado; Hillsborough County Public Schools, Florida; Montgomery County, Maryland; New Haven, Connecticut; Norfolk Public Schools; Plattsburgh City School District, New York; Poway Unified School District, California; Providence, Rhode Island; St. Francis Independent School District, Minnesota; San Juan Unified School District, California; and Toledo City School District, Ohio, among others.

They were able to replicate the collaborative approach to problem-solving used in negotiations in other joint labor-management committees. By changing the tone of their interactions from adversarial to collaborative, they were able to have conversations about the possibility of reform continuously and not solely in the context of contract negotiations. An important commonality among the case study districts was the commitment of district and union leaders to communicate with each other, formally and informally, to sustain a positive working relationship. Other studies have also emphasized that leaders must be willing and committed to working collaboratively if reform efforts are to be successful (Donaldson, 2011; McLaughlin, 1990).

In a review of two California districts' efforts to establish PAR programs, Humphrey, Koppich, Bland, and Bosetti (2011) identified the importance of joint-governance and shared responsibility. The Poway Unified School District and San Juan Unified School District worked collaboratively with their teachers unions to institute effective PAR programs. PAR programs require teachers to take on the role of evaluating their peers, which is a departure from traditional teacher evaluation systems. Both districts established Governance Boards prior to this work, which are joint labor-management bodies that oversee the program and make recommendations to the district on individual teachers. The district and union would share the responsibility of making difficult decisions on teachers' employment status, as well as other high-stakes issues. The Governance Boards served as incubators for labor-management collaboration. As a result of this work, collaboration became a part of district culture and dictated their approach to future reform efforts.

It should be noted that none of the case study districts referenced in this section became paradigms of collaboration overnight. Sometimes the work was difficult, but consensus is that their efforts to transition away from traditional adversarial practices and find ways to collaborate on reform were worthwhile. This hypothesis is consistent with what I observed in my three case study districts.

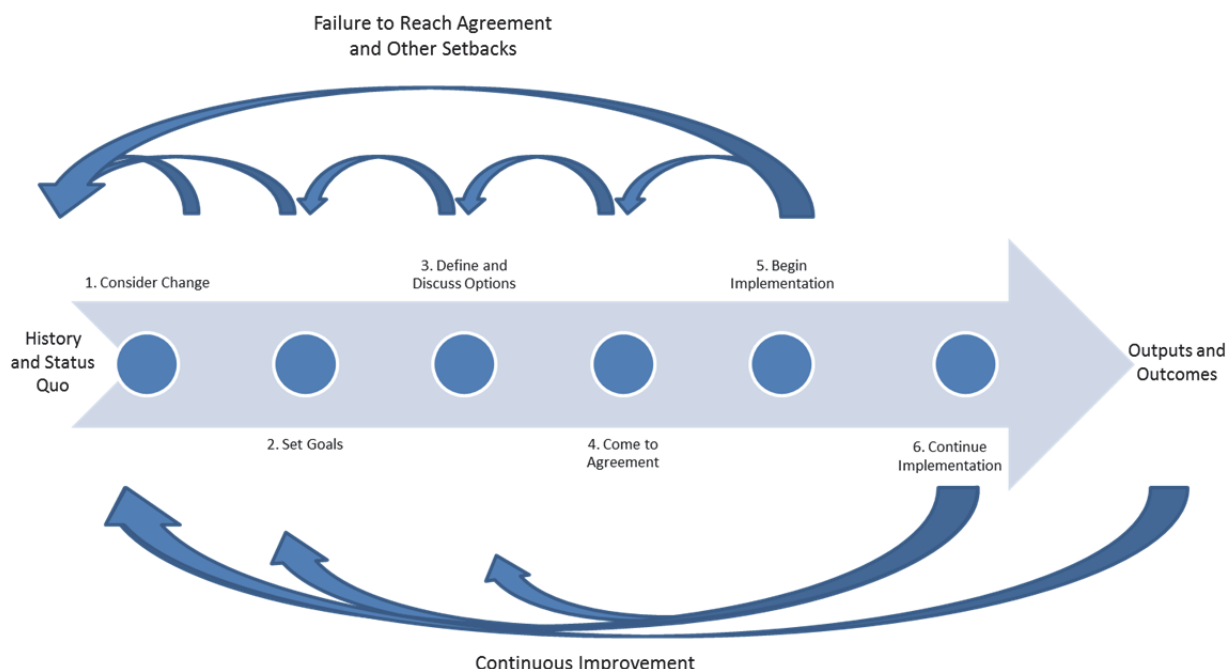
Chapter Four: Conceptual Framework

I developed a conceptual framework to guide my approach to this research. Given that the focus of my dissertation is the reform process itself, I created a model to outline the stages in that process. This framework informed my approach to collecting, organizing, and analyzing data. It also served as the canvas for telling the story of each case study district and for comparing cases. Several constructs describe the reform process, with the stages serving as the central organizing principle. I discuss this framework and related constructs below.

Reform Process Framework

The framework displayed in Figure 4.1 outlines a generic process through which education reforms are developed and implemented. I developed this model after familiarizing myself with various education reform efforts taken by school districts, and refined it after reviewing the data collected for my three case studies. It is designed to be broad enough to describe reforms in any area education policy, though I apply it exclusively to teacher evaluation reforms.

Figure 4.1 Reform Process Framework



This framework is linear; when one stage is complete, the next stage should begin. For example, in order for implementation to commence, stakeholders must agree on which policy option will be implemented. However, human decision-making is not always a linear process,

nor does it proceed uninterrupted. While this process is displayed as a linear trajectory, in reality, these efforts may not proceed linearly. Actors may encounter stumbling blocks where the process stalls or even regresses. Setbacks may result from the failure of key stakeholders to reach agreement, a loss of funding, changes to state laws, or other influential events, such as teacher strikes. The framework represents this regression with a series of arrows linking one position in the process to the preceding stages. These setbacks may vary in magnitude. An agreement to move forward with a particular policy option might fall through, which would then place the players in the position of re-defining and considering other policy alternatives. At the extreme end, the implementation of an adopted policy may fail or otherwise be abandoned, leaving stakeholders with a revised status quo that will need to be addressed with a new approach.

The reform process should also be viewed as a cycle. An evaluation of policy implementation and the subsequent outputs and outcomes provides information that motivates continuous improvement. This is displayed in Figure 4.1 with the arrows that link the final stages of the process, where the district is “moving forward” with the reform, back to earlier stages. For example, as implementation continues, stakeholders may identify flaws in their evaluation systems and decide to define and consider new policy options. My framework recognizes and allows for these feedback loops.

This process is composed of six stages, denoted by circles in the diagram, which trace reform efforts from the status quo that motivates stakeholders to act, to the eventual outputs and outcomes of the new policies that are implemented. The stakeholder interview protocols I developed followed this framework. I asked the interviewees to describe what transpired in their district at each step. Table 4.1 summarizes each stage, and groups the steps into four broader phases in the process. I later present my results in four chapters, one for each of these phases.

The bookends to this process are status quo of district policies and eventual outcomes of the new policy that will be implemented. While not a “stage,” history provides the context for my analysis of each district. The status quo, prior to the first stage of the reform process, provides insight into the district’s motivation for considering change. Therefore, I collected data on the history of district labor-management relationships, existing teacher evaluation policies, previous reform efforts, and other potentially motivating factors. On the other end of the reform process are the outputs and outcomes of newly implemented policies. Short-term outputs of my case study districts’ reform efforts will include the degree to which early implementation was successful, while long-term outcomes will include the effects of the policy on teacher effectiveness and student achievement. Though it is too early to collect data on policy outcomes in my case study districts, understanding the desired effects of their teacher evaluation reforms is important, for it is the potential of improved outcomes that motivates reform in the first place.

Table 4.1 Teacher Evaluation System Reform Process, by Reform Stage and Key Activities

| | Stage | Key Activities |
|----------------------------------|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Setting the Stage | 1) Consider change | <ul style="list-style-type: none"> • Initiate a conversation between district and union leaders • Consult with colleagues • Agree to form a teacher evaluation committee • Recruit committee members |
| Approach to Reform | 2) Set goals | <ul style="list-style-type: none"> • Organize the committee • Diagnose existing evaluation system • Explore policies in other districts • Discuss goals and concerns • Develop common understanding • Communicate with stakeholders |
| Hammering Out the Details | 3) Define and discuss options | <ul style="list-style-type: none"> • Research teacher evaluation policies • Identify policy options • Evaluate alternatives • Create materials for teacher evaluation pilot |
| | 4) Come to agreement | <ul style="list-style-type: none"> • Review policy documents • Negotiate terms of agreement/MOU • Agree to begin pilot |
| Moving Forward | 5) Begin implementation | <ul style="list-style-type: none"> • Recruit pilot participants • Train stakeholders • Monitor pilot implementation • Propose modifications |
| | 6) Implementation continues | <ul style="list-style-type: none"> • Continue pilot • Negotiate permanent agreement • Prepare for full implementation • Implement district-wide |

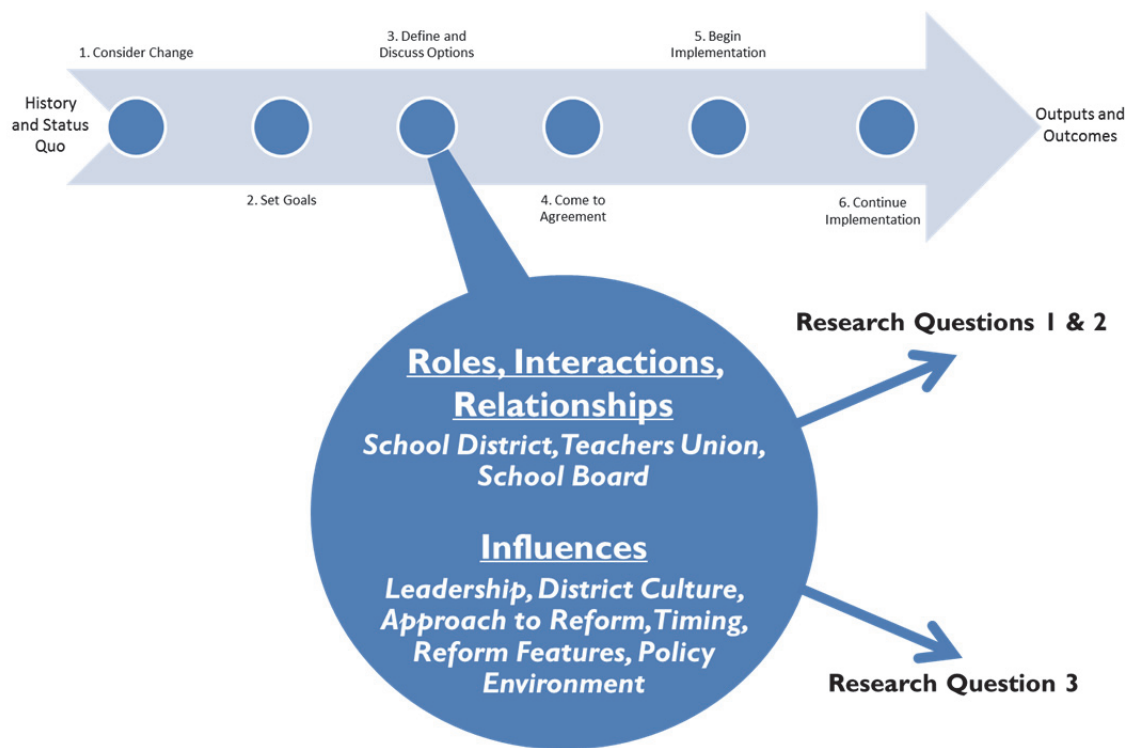
The first stage of the reform process, “consider change,” marks the point at which stakeholders agree to formally begin the reform process. The second stage, “set goals,” is when committees are convened and may include the drafting of formal goal statements and establishing norms. The next stage begins when stakeholders identify a set of policy options and consider their merits, and thus is labeled, “define and discuss options.” In the fourth stage, “come to agreement,” an official decision on which policy option will be adopted is made and the parties agree to move forward. The fifth stage is the first stage of implementation. This stage may begin with a pilot of the new policy in a few schools or with selected teachers. As implementation continues in the sixth stage, there may be a second year of the pilot and eventually implementation will spread to the entire district. The final stage does not “end” until a new policy is adopted.

Each stage in the reform process is defined by the actions of the key players – district administrators, teachers unions, and school board members – as well as a number of external factors. The stakeholder groups who are most central to the teacher evaluation reform process are the school districts and teachers unions.⁶ Their actions and interactions drive the process and influence its outcomes, and their relationships are unlikely to remain constant over time. The key

⁶ While school board members do shape district priorities and hire superintendents, I found that they are not as involved in designing teacher evaluation policies.

stakeholders' orientation to change and willingness compromise shape those interactions, which subsequently influence their ability to agree on changes to policy. In turn, the implementation of new policies can shape labor-management relations. District policy decisions are constrained by existing CBAs, state laws, federal laws, court rulings, school board policies, and the political influence of stakeholder groups, including teachers unions (Hess and Loup, 2008). My first two research questions focus on the key players, while my third looks at the range of influential factors, both internal and external. These factors may include the characteristics of district and union leadership, the approach they take to reform, the characteristics of the policies considered or adopted, and the timing of those efforts. To answer these questions, I describe the characteristics of the key players, as well as other internal and external factors within the context of each stage in the reform process, as displayed in Figure 4.2.

Figure 4.2 Within-Stage Analysis



In this report, I present my results stage-by-stage and then draw cross-cutting conclusions. However, I am only able to analyze the stages for which I have data. My case study districts have not completed the entire process outlined in my reform process framework. While two districts are currently implementing pilots, one has not moved beyond the second stage of setting goals.

Key Constructs

Several important constructs are identified in my research questions. Table 4.2 links each research question with its corresponding constructs. Since I took an exploratory approach to this research, I started with working definitions for a set of constructs, informed by the literature, and later refined them as I explored my data. As previously noted, the underlying principle that spans all three questions is the reform process, which is divided into six stages. In my analysis, I describe the constructs at each of the stages I observe.

Table 4.2 Key Constructs by Research Question

| Research Question | Key Constructs |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------|
| Central Organizing Principle | |
| Reform process | Stages |
| Research Question 1 | |
| What roles do teachers unions and districts play in efforts to reform teacher evaluations and to what extent does the involvement of these groups shift throughout the reform process? | Roles Involvement |
| Research Question 2 | |
| How do teachers unions and districts interact and how do their relationships change throughout the reform process? | Interactions Relationships |
| Research Question 3 | |
| What influences the roles, interactions, and relationships between teachers unions and districts during the reform process and the policy outcomes of their efforts? | Influential factors Degree of influence Direction of influence |

Roles in the Reform Process

The constructs related to my first research question are stakeholder “roles” and “involvement.” Stakeholders may play different roles and their involvement may vary throughout the reform process. I identified three main categories of participation in the reform process: leader, participant, and observer. To further differentiate individual involvement in reform efforts, I defined four levels of participation, from significant to superficial. (See Table 4.3.) To be a “leader,” one must be significantly or consistently involved at a given stage. A “participant” could be consistently or occasionally involved. A person who is superficially involved would fall into the category of “observer.” In addition to identifying the individuals involved, I note if there were key stakeholders or stakeholder groups who were not involved and possibly not even aware that the reform efforts in their districts were taking place.

Specific roles are defined by level of involvement and also the responsibilities individuals assume. There are multiple ways individuals can lead, participate, or observe. Based on my observations of the reform activities in each of the case study districts, I have identified ten specific roles within those three main categories. They are described in Table 4.4. While this is

not a comprehensive list of all the types of roles and contributions individuals can make to the reform process, they capture the majority of behaviors I observed in the data.

Table 4.3 Levels of Involvement

| Level of involvement | Description |
|----------------------|-------------------------------------------------------------------------------------------------------------------------------|
| Significant | A key player; the process would likely not have moved forward at a particular stage without that person. |
| Consistent | Involved regularly and makes meaningful contributions to the process. |
| Occasional | May participate from time-to-time, check-in on the process, or advise the other individuals. |
| Superficial | Pays attention and may support or criticize reform efforts, but does not participate in committees, pilots, or key decisions. |

Table 4.4 Description of Stakeholder Roles⁷

| Participation Category | Role | Description and Example Responsibilities |
|------------------------|------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Leader | Initiator | Responsible for getting the reform initiative or initial conversation started. Takes a proactive role. |
| | Convener | Brings people together, schedules the initial committee meeting, or recruits committee members. |
| | Facilitator | Runs meetings, serves as a chair or co-chair, leads discussions, may create agendas or bring materials, and helps keeps the process moving forward. |
| | Implementation Lead | Primarily responsible for policy implementation. Oversees the pilot evaluation, conducts trainings, monitors implementation, and supports participants. |
| Participant | Contributor | Committee member who actively contributes and completes tasks (e.g., doing research or note-taking). An active participant in the pilot program. |
| | Communicator | Provides updates on the reform efforts to the district community. Solicits feedback from stakeholders. |
| | Advisor | May occasionally attend meetings. Listens to discussions, expresses support, sets expectations, provides feedback or helps guide the process occasionally. |
| | Implementation Support | Support implementation. Materials, trainings, soliciting feedback from pilot participants. Assist with trainings. |
| Observer | Supporter | Pays attention to updates on the district's work. Believes the work is a priority. Expresses support or approval. |
| | Critic | Criticizes or disagrees with policy decisions made or does not believe reforms should be undertaken. |

There are four leadership roles, and they are relevant at different stages in the reform process. Initiators take on a leadership role in the first reform stage, “considering change.” Conveners

⁷ These categories describe the roles individuals take on during the teacher evaluation reform process, not the roles they may play in their respective organizations (i.e. the teachers union, district office, or school board).

also help get the process moving in the early stages. Once committees are formed, the facilitator role becomes necessary, and when the pilot begins, an implementation lead is needed. I identified four relevant roles played by participants who were not necessarily leaders. The broadest category is “contributor,” which includes individuals who served on teacher evaluation committees, participated in the pilot teacher evaluation – including teachers and school administrators – or who otherwise meaningfully contributed to the district’s reform efforts.

These roles are not mutually exclusive, as individuals can wear multiple hats at once or take on different roles at various stages in the reform process. While people often point to one overall “leader,” multiple individuals can serve in a leadership capacity. In the three case study districts, more than one individual took on a leadership role during the reform process. Additionally, some individuals may take on responsibilities that do not fit neatly into any one category. However, these categories capture the most common roles observed in the data.

The role of “observer” is not an active role, and signifies superficial involvement in the reform process. Still, it is important to recognize which stakeholders support, or perhaps oppose, the district’s efforts from the periphery. Someone may be involved by speaking out against the initiative and possibly discouraging others from supporting or participating in a pilot program. It is also possible for an active participant to object to the reform initiative itself and resist change, even while serving on a committee or participating in a pilot.

Labor-Management Relationships

As noted in Table 4.2, there are two constructs addressed by my second research question: interactions and relationships. Several typologies that describe labor-management relationships have been proposed in the literature and were discussed in Chapter 3. Walton and McKersie (1991) identified five labor-management relationship patterns: conflict, containment-aggression, accommodation, cooperation, and collusion. Susan Moore Johnson (1984) used three categories: collaborative, cooperative, and contentious. After consulting the literature, I chose three categories that best described the range of relationship characteristics observed in my data: contentious, civil, and collaborative. I define my three relationship categories by the level of trust between the parties, their interactions, and common behaviors of both union and district leaders. These are described in Table 4.5. I incorporated descriptors from the literature and from my own observations. This typology captures a continuous range of relationship characteristics, rather than a discrete set of three relationship types. Relationships are fluid and can move across or settle in between these categories. Further, the interactions observed may match only some of the descriptors proposed in Table 4.5. None of these relationship types imply that the two parties always agree or disagree. Collaborators can engage in hard conversations and respectfully disagree with each other. It is also possible for adversaries to come to agreement on various issues. Another important dimension of these relationships is whether or not they can be sustained over time. Thus, I will note if stakeholders are able to maintain cooperative

relationships at multiple stages in the process, or if early efforts to collaborate later devolve into conflict.

Table 4.5 Characteristics of Labor-Management Relationships

| Relationship Type | Level of Trust⁸ | Typical interactions | Union behaviors | District leader behaviors |
|--------------------------|-----------------------------------|---------------------------------------------------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| Contentious | Distrust | Combative, adversarial, uncooperative, antagonistic, formal | Reactive, obstructive, solidarity/ mobilize membership, competitive, oppositional | Authoritarian, hostile, rigid, competitive |
| Civil | Limited | Courteous, somewhat cooperative, accommodating | Internally-focused, hands-off, giving-in | Focus on compliance, considerate |
| Collaborative | Extended | Respectful, friendly, cooperative, diplomatic, conciliatory, polite | Proactive, open to compromise | Reach out to union leaders, open to compromise, focus on problem-solving |

The labor-management relationship categories utilized in this study relate to the type of collective bargaining and paradigms of unionism in each district and align with the interactions observed in previous studies. Districts may engage in traditional collective bargaining or in IBB, and the latter is more aligned with the goals of reform unionism. Interactions during contract negotiations, while not limited to the topic of teacher evaluation, are relevant to this research, as they may reflect or impact district-union relationships. As previously discussed, Gates and Vesneske (2011) described interactions during contract negotiations with five action frames: craftsmanship, stewardship, competition, solidarity, and diplomacy. They derived these categories by observing the behaviors of district and union leaders. Their definition of competition, where parties focus on winning, rather than compromising, aligns with the contentious relationship category adopted for my study. On the other hand, the diplomacy action frame described efforts to solve problems and compromise, which are behaviors that match my description of collaborative relationships, and also are indicative of the IBB approach to collective bargaining. I also observed behaviors that aligned with Gates and Vesneske's solidarity approach, where union leaders actively solicit support from and mobilize their membership against a proposed policy change advocated for by the district, and use that resistance as a bargaining tool. This union behavior is linked with the contentious relationship type. However, it should be noted that union leadership can also mobilize membership in support of reform efforts in cooperation with district leadership.

⁸ As defined by Walton and McKersie (1991).

Influential Factors

To answer my third research question, I addressed the many possible factors that influenced the reform process. These include any factor that significantly shaped stakeholders' roles, relationships and interactions; how their work progressed; and the policy outcomes of their efforts. The major factors fall into five broad categories:

1. Leadership;
2. Approach to reform;
3. Reform features;
4. District culture and characteristics; and
5. Timing and external events.

The first two categories are relatively straightforward: district and union leaders and other individuals involved in the reform efforts have a direct influence on the reform process and its outcomes. Therefore, the experience of and degree to which a leader is effective makes a difference. Turnover in leadership is also an important factor. Further, the relationships between the key players influence the way they interact with each other. Johnson (1984) found that individual personalities and relationships had an impact of reform efforts. Relationships are also shaped by prior interactions, and also by past labor-management conflict experienced in the district. In a way, relationships can be seen as an intermediate factor. Leaders and other factors affect the character of a relationship, which, in turn, can influence an outcome that further shapes that relationship.

“Approach to reform” addresses how the work was organized. This includes the structure of teacher evaluation committees, the planning process, and the approach to implementing pilot teacher evaluation programs. The degree to which reform efforts were a joint-approach between labor and management is relevant here. The third factor, reform features, includes not only the features of the policies that were ultimately adopted, but any that were discussed considered as policy options. This also includes “the stakes” attached to the teacher evaluation policies (i.e. the consequences of evaluation results).

The fourth factor, district characteristics, encompasses district organizational culture, historical events, and demographics. This also includes school culture, which may vary within a district. Negotiations around teacher evaluations may also be influenced by characteristics of the district's student population. Teachers in low-performing schools and districts, which may enroll more students from disadvantaged backgrounds, may have greater concerns about teacher evaluations. They may fear that they will be penalized for the performance of their students, which is likely influenced by demographic characteristics that are beyond their control. Likewise, leaders of low-performing districts may have the most incentive to promote these reforms in an effort to raise student achievement. In case studies of six districts, Johnson (1984) found that district size could explain some of the variation in labor practices, but that other factors, such as whether the district was urban, suburban, or rural, had little influence. On the

other hand, Johnson noted that several factors did not explain the variation observed across school districts. These included the number of years the district had been engaged in collective bargaining, urbanicity, and the number of previous teacher strikes.

The fifth factor refers to the influence of the timing of the reform efforts with other significant events. These include the arrival of new leaders, the implementation of new policies, such as the Common Core standards, and also the availability of financial resources as the state recovered from the Great Recession. This category includes external events that lie outside of districts' control, such as state policy decisions and media coverage.

There are two implied constructs that further define these factors: 1) direction of influence – whether a factor is a barrier or enabler; and 2) degree of influence – how much of an impact it had. The degree of influence can be minimal or great, and at the extreme end, can cause the process to stall. For my analysis, I identified influences in both directions – positive and negative – and focused on factors that had the greatest degree of influence.

Chapter Five: Research Design

This chapter presents the research methodology I used for my dissertation. My research questions, which are descriptive in nature, are best addressed through case studies. Therefore, my research followed a multiple case exploratory design. The cases consist of three school districts in California that have local teachers unions and began engaging in work to reform their teacher evaluation reforms between 2010 and 2011. I collected primary data through semi-structured interviews with key stakeholders, including school district leaders, teachers union leaders, and school board members. I also collected secondary data in the form as documents, such as CBAs, school board meeting minutes, and policy descriptions. My conceptual framework provided the structure for data collection and analysis. I applied qualitative codes to interview transcripts and other documents to identify important themes and patterns.

Case Study Districts

Case studies are a common way that researchers conduct analyses of activities in individual schools and districts. I selected three school districts purposively, according to a set of criteria. I selected districts that would provide unique, illustrative cases of the teacher evaluation reform process. Conducting case studies allowed me to illuminate details and establish strong internal validity within each case, while also describing the frequency and range of themes across cases. The qualitative data I collected enabled a deep exploration, though it did not provide broad generalizability. Thus, I used my small, strategically-selected sample to explore, in-depth, the particular circumstances and mechanisms that operated within each district.

I selected districts within California because the state has not made any recent policy changes regarding the composition of teacher evaluations. Additionally, California has not received federal Race to the Top funds or an NCLB waiver, both of which would have required the state to commit to specific teacher evaluation policy changes, such as mandating the inclusion of measures of student achievement as a significant factor in teacher effectiveness scores. As a result, efforts to reform teacher evaluations in California are largely motivated at the district level, which is the focus of this research. The incentives that motivate reforms in California are likely different than in other states and may include financial incentives, public pressure, or expectations that the content of teacher evaluations will be mandated by future legislation. Holding state context constant also controlled for the influence of collective bargaining legislation, which varies across states. Several other states have recently passed legislation to restrict or ban collective bargaining. In those states, the influence of teachers unions has been curtailed and the relationship between stakeholders at the district level is in a state of flux.

California School District Characteristics

As of the 2013-2014 school year, there were 949 school districts in California⁹⁹, which enrolled over 6.1 million students and employed over 300,000 teachers. Appendix A. includes characteristics for the entire population of school districts in the state, where data were available. California school districts vary greatly by size, from 653,826 students enrolled in the Los Angeles Unified School District, which is also the second largest school district in the country, to only 3 students in the Panoche Elementary School District in San Benito (California Department of Education, 2014). Table 5.1 shows the distribution of school districts in California by size, as well as the total number of students enrolled, by district size. The largest 10 school districts educate one-fifth of all state public school students and 85% of all public school students are enrolled in districts with at least 5,000 students.

Table 5.1 Distribution of School Districts in California By Enrollment

| Enrollment | Number of Districts | Percent of Districts | Number of students enrolled | Percent of all California students enrolled | Cumulative percent of students enrolled |
|------------------------|---------------------|----------------------|-----------------------------|---------------------------------------------|-----------------------------------------|
| 50,000+ | 10 | 1% | 1,277,655 | 21% | 21% |
| 40,000 - 50,000 | 7 | 1% | 314,467 | 5% | 26% |
| 30,000 - 40,000 | 16 | 2% | 531,307 | 9% | 35% |
| 20,000 - 30,000 | 41 | 4% | 987,880 | 16% | 51% |
| 10,000 - 20,000 | 90 | 9% | 1,270,885 | 21% | 71% |
| 5,000 - 10,000 | 115 | 12% | 848,251 | 14% | 85% |
| 1,000 - 5,000 | 296 | 31% | 779,719 | 13% | 98% |
| 0 - 1,000 | 374 | 39% | 124,645 | 2% | 100% |
| Total | 949 | 100% | 6,134,809 | 100% | |

Note: Data is accurate as of the 2013-2014 school year. Only traditional elementary, high school, or unified school districts are included. Totals may not add up to 100% due to rounding.

Source: California Department of Education, "Student & School Data Files," 2014. As of December 9, 2014: <http://www.cde.ca.gov/ds/sd/sd/index.asp>

Local teachers unions in California are affiliated with the California Teachers Association (CTA) or the California Federation of Teachers (CFT) (or both), which are affiliated with National Education Association (NEA) and the American Federation of Teachers (AFT), respectively. The CTA has over 1,300 local affiliates and the CFT has over 140 local branches across the state (cite CFT and CTA websites). Local affiliates vary greatly in size, as do the number of teachers districts employ. While the Los Angeles Unified School District employs over 32,000 teachers, more than 100 districts employ fewer than 10 teachers each. The distribution of districts by the number of teachers employed is shown in Table 5.2.

⁹⁹ This includes traditional elementary, high school, or unified school districts.

Table 5.2 Distribution of School Districts in California by Teachers Employed

| Number of Teachers | Number of Districts | Percent of Districts |
|----------------------|---------------------|----------------------|
| 3,000 | 5 | 0.5% |
| 2,000 – 3,000 | 9 | 0.9% |
| 1,000 – 2,000 | 39 | 4% |
| 500 – 1,000 | 100 | 11% |
| 250 – 500 | 111 | 12% |
| 0 – 250 | 684 | 72% |

Note: Only traditional elementary, high school, or unified school districts are included. Totals may not add up to 100% due to rounding. The most recent data available was for the 2011-2012 school year. Source: California Department of Education, "Student & School Data Files," 2014. As of December 9, 2014: <http://www.cde.ca.gov/ds/sd/sd/index.asp>

Selection Criteria

I chose districts with some contextual similarities, while also ensuring a degree of cross-case variation. State context is held constant, but the districts vary on a number of demographic characteristics, including size and student performance. I intended to include districts that demonstrate varying degrees of district-union collaboration, though this was difficult to assess prior to recruiting districts. These demographic differences enable me to describe and compare a range of experiences. My core criteria for selecting districts were:

1. Must be engaged in the process of reforming their teacher evaluation policies; and
2. Must have a teachers union and engage in collective bargaining.

I also considered enrollment when selecting case study districts. I wanted to capture variation across case districts, and size may have interesting implications for the nature of reform efforts. I defined four size categories to sort the population of school districts in California:

1. Large districts: at least 20,000 students
2. Medium districts: between 10,000 and 20,000 students
3. Small districts: between 5,000 and 10,000 students
4. Very small districts: fewer than 5,000 students

I decided to exclude very small districts from my sample for two key reasons. First, very small districts would have few stakeholders to interview and it would be difficult to verify reports of informants. It may take larger districts longer to negotiate with stakeholders and make policy changes, which would make the decision-making process more complex. Second, very small districts, with fewer resources and fewer teachers, may be less likely to independently tackle teacher evaluation reform. There are a total of 279 school districts in California that met this size criterion as of the 2013-2014 school year. Collectively, these districts educate 85% of the students in California public schools. I also wanted to capture variation in district size, to explore whether or not it emerges as an important influential factor. Therefore, I selected one small, one medium, and one large district for my case studies.

District Sample and Recruitment

To identify candidate districts, I conducted an internet search of district websites and news articles. I searched for documents related to teacher evaluation – district newsletters, contracts, teachers union publications, school board meeting agendas and minutes, and local news articles. This search yielded eight candidate districts. I began recruiting districts in October 2013 and secured three participants by April 2014. In the first stage of recruitment, I reached out to five districts. I emailed district superintendents and also contacted district research offices (where they existed) by email and phone. When I did not receive a response, I followed-up several times. I was able to submit formal research applications in three districts, all of which, after a lengthy review period, were denied, despite encouraging conversations with the research offices. The superintendent from one of the initial districts I contacted did reply and agreed to participate; no formal research application process was required. I then contacted the remaining three districts and ultimately received the permission of two superintendents to conduct my research, securing the third district in April 2014.

To aid in the comparison of my case study districts, I defined four categories to describe six demographic characteristics: enrollment, English Language Learner (ELL) population, number of students qualifying for free or reduced price lunch under the National School Lunch Program (FRPL), minority population, schools meeting their targets on California's Academic Performance Index (API), and poverty level in the surrounding city. (See Table 5.3) The categories were informed by the distribution of school districts across the state for each characteristic. These categories allow me to describe case study districts in this report without using precise information that would reveal their identities. I also assigned pseudonyms (Laurel, Juniper, and Sumac) to each district and genders to different stakeholder roles¹⁰, to maintain anonymity. Table 5.4 describes the three case study districts by the demographic classifications.

¹⁰ In my writing, I refer to superintendents as female, other district and school administrators as male, teachers union leaders and other teachers as female, and school board members as male.

Table 5.3 Descriptors for Select School District Characteristics

| Demographic | Descriptors | | | |
|--------------------------------------|-------------|-----------------|---------------|-----------------|
| | Large | Medium | Small | Very Small |
| Enrollment | ≥ 20,000 | 10,000 – 19,999 | 5,000 – 9,999 | < 5,000 |
| | High | Medium | Low | Very Low |
| % ELL | ≥ 30% | 15% - 29% | 5% - 14% | < 5% |
| % FRPL | ≥ 75% | 50% - 74% | 25% - 49% | < 25% |
| % Minority | ≥ 75% | 50% - 74% | 25% - 49% | < 25% |
| % Schools meeting API targets | ≥ 75% | 50% - 74% | 25% - 49% | < 25% |
| Poverty (city level) | ≥ 15% | 10% - 14% | 5% - 9% | < 5% |

Note: Data for school-level indicators are accurate as of the 2013-2014 school year, except for the API target data, which is for the 2012-2013 school year.

Sources: California Department of Education, "Student & School Data Files," 2014. As of December 9, 2014: <http://www.cde.ca.gov/ds/sd/sd/index.asp>; U.S. Census Bureau, "State and County QuickFacts," 2014. As of December 9, 2014: <http://www.census.gov/quickfacts>.

Table 5.4 District Sample Characteristics

| Indicator | Juniper (JUSD) | Laurel (LUSD) | Sumac (SUSD) |
|--------------------------------------|----------------|---------------|--------------|
| Size | Small | Medium | Large |
| % ELL | Very Low | Low | Medium |
| % FRPL | Very Low | Low | High |
| % Minority | Low | Medium* | High |
| % Schools meeting API targets | High | Medium | Very Low |
| % Poverty (city level) | Very Low | Medium | High |

* Laurel varied on the border of "low" and "medium" percent minority for the past few years. Given that their minority population was consistently higher than Juniper's, I decided to place Laurel in the "medium" category for minority population.

There are some clear distinctions between the three districts based on demographics alone. Juniper is the smallest of the group and also the highest-performing. It is located in the wealthiest of the three cities and has the least diverse student population. On the other end of the spectrum, Sumac is the largest of the group and also the lowest-performing. It is the most diverse of the three districts, with the largest populations of minority students, English Language Learners, and students who qualify for free or reduced-price lunch. Laurel falls between the other two districts in all six demographic categories. It is medium-sized and, compared to the rest of the state, has small populations of English Language Learners and students who qualify for free or reduced-price lunch. The majority of its schools met their API targets in the 2013-2014 school year.

Document Collection

I collected documents to provide background and context for each of the case study districts and to fill in information that may have been missing from stakeholder interviews. Some documents were provided to me by the study participants. The documents I collected included:

- Current and past labor contracts
- MOUs related to teacher evaluation policies
- School board meeting minutes and links to videos
- District teacher evaluation policy documents
- Vision/goal statements
- Teacher evaluation committee agendas, notes, and final products
- Research and resources utilized by committees
- State policy documents
- District or union newsletters
- Presentations
- News articles

These documents served as artifacts to corroborate details about past events in the districts. Since interviewees may have imperfect memories of past events, these documents strengthened my understanding the beginning stages of district reform efforts, as well as past reform efforts.

Stakeholder Interviews

The primary data source for this study is interviews with individual stakeholders, including representatives from school district offices, teachers unions, and local school boards. I interviewed 8-11 individuals from each district. Interviews were conducted in-person when possible, and over the phone when in-person meetings could not be arranged. Each interview lasted approximately one hour. I began recruiting participants and conducting interviews once I received district approval to conduct my research. I conducted the majority of interviews from January 2014 through June 2014. I interviewed two additional stakeholders in December 2014. I also conducted three follow-up interviews with the key district representatives from each district in December 2014, for a total of 32 interviews.

Stakeholder Recruitment and Selection

Interviewees were purposively selected to ensure representation from the three key stakeholder groups: district administrators, teachers union representatives, and school board members. I first identified individuals in leadership positions and then applied a snowball method to identify additional contacts based on recommendations from past interviewees. In some cases, a district administrator would email the recommended individuals to first provide an introduction and I would follow-up if they agreed to participate. I continued to recruit participants until I had at least eight participants per district and representatives from all three

stakeholder groups. While I did not initially recruit any school administrators, a district administrator in Sumac chose to invite two of them to participate and they agreed. In total, I contacted a total of 38 individuals, 29 of whom agreed to participate in an interview. The email used to recruit participants is included in Appendix B. My final sample of 29 interviewees is described in Table 5.5.

Table 5.5 Stakeholders Interviewed by District

| | Juniper | Laurel | Sumac |
|-----------------------------------------------------------|----------------|---------------|--------------|
| Superintendent | 1 | 1 | 1 |
| Assistant Superintendent of Human Resources ¹¹ | 1 | 1 | 1 |
| Other District Administrators | - | 2 | 3 |
| School Administrators | | | 2 |
| Teachers Union Presidents ¹² | 2* | 1 | 1 |
| Teachers Union Representatives ¹³ | 2 | 3 | 2 |
| School Board Members | 2 | 2 | 1 |
| Total | 8 | 10 | 11 |

*Teachers union leadership changed over the course of the teacher evaluation work. In Juniper, I interviewed the individual who was president when the first teacher evaluation committee was formed (UP #1), as well as the current union president (UP #3), who was in office when the third teacher evaluation committee was formed.

Interview Protocol

I conducted semi-structured interviews with the selected stakeholders from each district; one individual from each district was interviewed twice to provide updates. This approach allowed me to address the same topics across all interviewees, while providing a degree of flexibility so that the participant could lead the conversation and not feel limited by a rigid set of questions. Given that the primary aim of this research is to characterize the reform process itself, the steps in the reform process, as outlined in my conceptual framework, served as the central organizing principle of the interview protocol. The full interview protocol is included in Appendix C.

¹¹ To protect participant anonymity and maintain consistency, I refer to the most senior individual responsible for human resources as the “Assistant Superintendent of Human Resources,” or “Assistant Superintendent of HR.”

¹² Juniper had three different union presidents during the teacher evaluation reform efforts. I will refer to them at “UP #1,” “UP #2,” and “UP #3.”

¹³ These representatives may not be elected or hold a title, but may be committee members representing teachers or participating in evaluation pilots.

The interview was divided into the following five sections:

1. Background Questions
2. Labor-Management Relationships
3. Reform Process Overview
4. Roles, Interactions, and Relationships
5. Concluding Questions

The background questions established the interviewee's relevant prior experience and roles and responsibilities in his or her current position within the local education system. The second section solicited interviewees' perceptions of the labor-management relations in the district, in general, and also during the work to change teacher evaluation policies.

The third section was less structured, which gave the interviewee the opportunity to describe the reform process from his or her unique perspective, thus limiting the influence of my preconceptions as the interviewer. The grand tour question that was provided as a prompt to the interviewee was: "In your own words, could you please walk me through the efforts to change teacher evaluation policies in your district and how that work has evolved?" I guided the ensuing discussion with probing questions about the motivation for and goals of their work, as well as the reform features they considered and current state of the reform efforts.

In the fourth section, I asked the interviewee to review and elaborate on the information provided in his or her overview of the reform process. I used several probing questions to elicit greater detail and address any significant holes in the data. These probing questions addressed the key constructs included in my research questions. I asked stakeholders to describe their own roles in the teacher evaluation reform process, as well as the roles of others; comment on labor-management interactions; and provide their perspectives on what factors shaped the character of the reform process.

The interviews concluded with a few wrap-up questions, giving the interviewee the opportunity to provide any final thoughts or insights. I asked interviewees to share any relevant documents and suggest other stakeholders who could provide additional information. I took hand-written notes during the interviews and also audio recorded the conversations. After each interview, I summarized my notes and later used a transcription service to transcribe the dialogue from all the interviews.

Analytic Approach

After collecting both primary and secondary data, I took an exploratory approach to answering my three research questions:

1. What roles do teachers unions and districts play in efforts to reform teacher evaluations and to what extent does the involvement of these groups shift throughout the reform process?

2. How do teachers unions and districts interact and how do their relationships change throughout the reform process?
3. What influences the roles, interactions, and relationships between teachers unions and districts during the reform process and the policy outcomes of their efforts?

Since my data are in the form of written texts – interview notes and other documents – I applied qualitative analysis techniques. My analysis proceeded in four steps:

1. Construct case descriptions;
2. Apply structural codes to the text;
3. Identify cross-case themes; and
4. Analyze the themes.

My case study districts had not yet fully implemented new teacher evaluations during my data collection phase. Though I observed early implementation in Juniper and Sumac, I did not capture rich data in the final stage of the reform process. Additionally, Laurel did not move on from the goal-setting stage. Therefore, my analysis focused on the decision-making stages that precede implementation, as well as the beginning of the implementation phase.

Step 1: Construct Case Descriptions

I summarized the story of each district to highlight the within-case reform mechanisms. I pieced together the timeline for each district's reform efforts by comparing my notes from the various stakeholder interviews and cross-referencing those notes with other documents. I created a written profile for each district, informed by qualitative data from interviews and other records. The profiles include demographic information, characteristics of the existing and proposed evaluation systems, a description of the labor-management relationships, a directory of the stakeholder groups involved, and historical information regarding past reform efforts. I have integrated this information into my report.

Step 2: Apply Structural Codes to the Text

Interview transcripts and other documents were uploaded into the coding software QSR NVivo 9. I first labeled each document with the name of the district and further categorized interview transcripts with the date of the interview, stakeholder category (district administrator, union representative, school board member, or school administrator), the interviewee's job title, and whether or not they participated in their district's teacher evaluation committee. These categories allowed me to sort responses by district and type of stakeholder.

I read through the transcripts and applied a set of structural codes or "nodes," which are essentially a priori themes that match the structure of the interview protocol (Ryan and Bernard, 2003). This high-level coding allowed me to organize interviewee's responses across two dimensions. (See Table 5.6.) Process codes matched the stages of the reform process, as defined in the conceptual framework. The reform characteristics codes were drawn from the constructs identified in the research questions. These two sets of codes are not mutually exclusive; in

general, any quote describing a reform characteristic would also be associated with a reform stage. This coding scheme allowed me to align the stakeholder responses within each case and also facilitated comparisons across cases.

Table 5.6 Structural Codes

| Reform Process (Stages) | Reform Characteristics |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> • History and Status Quo • Consider Change • Set Goals • Define and Discuss Options • Come to Agreement • Begin Implementation • Continue Implementation • Current State of Reform • Next Steps • Quality Improvement Efforts • Setbacks | <ul style="list-style-type: none"> • Reform Features • Committee Characteristics • Stakeholder Roles • Labor-Management Interactions & Relationships <ul style="list-style-type: none"> ○ Positive ○ Negative ○ Negotiations • Influences <ul style="list-style-type: none"> ○ Barrier ○ Enabler |

In addition to the codes for the six reform stages, I applied process codes to categorize text related to history and status quo, the current state of reform in each district, quality improvement efforts in districts that began implementation, and setbacks that were experienced. I used the reform characteristics codes to isolate descriptions of stakeholder roles and relationships, teacher evaluation policies and committees, as well as influential factors in the reform process. I further differentiated some of the codes with sub-codes, which are described in Table 5.7. As I coded, I continually updated the district timelines and summaries, and consulted documents referenced in the interview transcripts.

Table 5.7 Reform Characteristics Code Descriptions

| Characteristic | Description | Example |
|-----------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Reform Features | Any reference to features of teacher evaluation policies, such as observations, ratings, and student achievement data. May refer to existing policies, topics of discussion, or features of teacher evaluation pilots. Includes opinions about related topics. | <i>"Well, I think the idea of peer review was... As we looked at different evaluation systems, I think we saw that that could really be used here."</i> |
| Committee Characteristics | Descriptions of teacher evaluation committees, including membership, group norms, and frequency of meetings. | <i>"The new committee was formed for this school year based on our agreement in negotiations... and it basically consists of four, sorry, three district, four union, one board observer; consensus is a vote of five on any item to change."</i> |
| Stakeholder Roles | Descriptions of the roles played by any stakeholder at any point in the reform process. Includes references to the level of involvement in the reform process. | <i>"I think I was just there more to listen and to give personal experience and perspective and to let the task force know that in leadership, we had an interest in making this both a more inclusive and productive process."</i> |
| Labor-Management Interactions & Relationships | Any reference to interactions between the district and teachers union or descriptions of their relationships. | |
| | <i>Positive</i> Denotes descriptions of positive interactions and relationships. | <i>"There's a lot of things that we're agreeing on, so I think the climate of the committee is definitely collaborative. I think it's moving in a positive direction."</i> |
| | <i>Negative</i> Denotes descriptions of negative interactions and relationships. | <i>"Well, the labor management relations at our district are probably at the absolute lowest that they have been in decades."</i> |
| | <i>Negotiations</i> Denotes text describing formal contract negotiations that may or may not be related to teacher evaluation. | <i>The bargaining team that was in negotiations was aware of this, because even at one time, as a committee, we worked up, "What can we share back with the bargaining team?" And it was really so that the bargaining team knew that this was on the radar.</i> |
| Influences | Any reference to a factor that influenced a district's progress or the labor-management interactions and relationships. Includes proposed explanations for why districts were or were not successful and why interactions were positive or negative. | |
| | <i>Barrier</i> Denotes descriptions of barriers to progress. | <i>"I think the down side of that is, what happens many times is, we do very substantive, comprehensive, highly collaborative processes that yield very little results. Because by the time you're done you're so exhausted because it took so much out of you."</i> |
| | <i>Enabler</i> Denotes descriptions of enablers of progress. | <i>"I think it's the process that we used. It wasn't top-down. Let's come together, figure out what our interests are, and see if this can serve both the district and the teachers, and we found that common ground. I think that's why it worked."</i> |

Step 3: Identify Cross-Case Themes

While my understanding of the key constructs in my conceptual framework was informed by prior research, I did not adopt a discrete set of constructs prior to collecting my data. Doing so

may have constrained or biased my analysis. Instead, I adopted exploratory approach and defined a set of explanatory themes while reviewing the data. I used the pile sorting technique to identify themes that describe the variation seen across cases as they emerged in the stakeholder interviews (Lincoln and Guba, 1985). This process ensured that I would draw conclusions authentically from the data and not solely from my own preconceptions.

I began by sorting the interview text according to the process codes, isolating the references to each stage. I then reviewed the references to the reform characteristics codes within each stage and across all three case study districts, grouping quotes according to emerging themes. After exploring the themes in each reform stage, I developed a coding tree by linking the themes identified to the related characteristics (Miles and Huberman, 1994). I defined a number of sub-codes for several of the characteristics codes, which were previously described in the conceptual framework chapter. For example, after exploring the descriptions of stakeholder roles within each stage, I identified the most common roles played by the various stakeholders. I adopted three broad categories – leader, participant, and observer – and described ten roles that I used to discuss the reform activities in all districts. This process helped me define the constructs embedded in my research questions. It would have been impossible to write a comprehensive list of what those roles were prior to collecting and analyzing the data. I used this same approach to characterize stakeholder relationships, settling on three over-arching relationship types: collaborative, civil, and contentious. Finally, I reviewed the intersection of the “influences” codes and process codes to identify a number of salient factors that influenced the reform process, including leadership, district culture, and approach to reform.

Step 4: Analyze the Themes

Cross-case comparisons contributed to my understanding of the reform mechanisms within each district and also provided insight into over-arching trends. After reviewing the interview transcripts and defining my explanatory themes, I analyzed their range and frequency. If the same trends were seen at multiple sites, then I used the data to build explanations for why those trends were prevalent. Likewise, when there was variation, I used the data to explain why there were significant differences across cases. I looked for the most and least typical themes, the most central themes, and the co-occurrence of themes across reform stages and districts (Ryan and Bernard, 2003).

This analysis illuminated several factors that had positive or negative effects on stakeholder interactions. Interestingly, one theme, such as leadership, can serve as an enabler to the reform efforts in one district, while posing a barrier to progress in another. I compared the structure and activities of each district’s teacher evaluation committees to understand what arrangements were most effective. I describe the strategies that contributed to success in individual districts, keeping in mind the unique characteristics of each district. Strategies that are effective in one district will not necessarily produce the same results in another district. Thus I tried to identify the aspects of district culture that influenced reform efforts and labor-management relationships. At the same

time I looked for concerns about teacher evaluations shared by stakeholders across districts, which were often the source of conflict. I present the results of this analysis for the six reform stages in four chapters, grouped into four large phases:

1. Setting the Stage (History and Status Quo; Consider Change)
2. Approach to Reform (Set Goals)
3. Hammering Out the Details (Define and Discuss Options; Come to Agreement)
4. Moving Forward (Begin Implementation; Continue Implementation)

Using the results of this analysis, I addressed each of my three research questions and proposed recommendations for policy and further research.

Limitations

The main limitation of this dissertation is the extent to which the results are generalizable beyond the cases studied. There is a limit to the external validity of this study; it cannot be assumed that the inferences made in this report will apply to all school districts. The number of stakeholders I interviewed affected the internal validity of each case study. Another limitation of the interviews is the possibility that responses were colored by individual bias or the desire to convey a certain message or promote an agenda. This was addressed by including a range of stakeholders – not only representatives from the school district or teachers union – and by comparing the perspectives of all the stakeholders within each case. To clarify areas where interviewees' responses did not cohere or where there was insufficient data, I cross-referenced the primary data with the documents I collected. For example, I was able to mine school board meeting minutes to confirm the dates of certain events and look at CBAs to clarify district policies. I also addressed areas that needed clarification with the follow-up interviews I conducted in each district.

The timeframe of my study also limited the volume of data I collected. Districts had already completed some stages prior to my interviews, and there will be reform activities that occurred after I conducted the stakeholder interviews. My analysis is limited to the stages for which I have data. One district, Laurel, only reached the goal-setting stage and did not progress to the later stages. Thus, my understanding of the subsequent stages is informed by only two districts, instead of three. I also relied on retrospective data to describe the earlier stages that the districts completed prior to my interviews. Again, the documents I collected helped me to verify this retrospective data and fill in the gaps where the interviewees' memories were imperfect. Despite the limits to generalizability and potential challenges to establishing internal validity, the lessons learned from these case studies will be of value to other districts engaging in or interested in pursuing similar work, policymakers drafting and considering legislation, and researchers studying related policies.

Chapter Six: Case Study District Overview

In this chapter, I provide an overview of each district's reform efforts. I first compare the characteristics of their pre-existing teacher evaluation policies. I then discuss the status quo of their labor-management relationships and present a timeline for the teacher evaluation reform process in each district. (Full case summaries are included in Appendices D, E, and F.)

Existing Teacher Evaluation Policies

While teacher evaluation policies are receiving renewed attention in the education community, all three case study districts had previously tackled this policy issue. The Stull Act lays out the basic requirements for teacher evaluation in California; however, the procedures for evaluating teachers are negotiated in locally-bargained contracts. In this section, I review the basic requirements of the Stull Act, and then note some of the details of each district's pre-existing evaluation policies. Table 6.1 describes key features of each district's teacher evaluation policy, prior to the initiation of current reform efforts in 2010 or 2011.

California Requirements

The frequency with which evaluations are conducted depends on a teacher's professional status. As required by the Stull Act, evaluations must be conducted annually for probationary teachers, and then once every two to five years for permanent teachers. Should a permanent teacher receive a rating of less than "satisfactory" in any area, an evaluation must be conducted again the following year until a "satisfactory" rating is achieved. During each evaluation cycle, districts must conduct at least one formal classroom observation of a teacher and have an end-of-year meeting. Districts are also required by the Stull Act to provide professional growth opportunities to teachers who do not reach the "satisfactory" level. This requirement addresses one of the primary goals of an evaluation system, which is to motivate improvement in teachers' practice. Administrators may choose to conduct more frequent observations of teachers who are struggling.

California law requires that districts evaluate teachers across four areas:

- The progress of pupils toward standards of expected pupil achievement;
- Instructional techniques and strategies;
- Adherence to curricular objectives; and
- Establishment and maintenance of a suitable learning environment.¹⁴

¹⁴ See California Education Code § 44662.

While the “progress of pupils” must be evaluated, California does not mandate that student achievement data be used to determine a teacher’s overall rating, either in whole or in part.

In addition to the four areas listed above, state law allows districts to incorporate any of the National Board for Professional Teaching Standards (NBPTS) or the California Standards for the Teaching Profession (CSTP). The CSTPs include six over-arching standards (California Commission on Teacher Credentialing, 2009):

1. Engaging and supporting all students in learning.
2. Creating and maintaining effective environments for student learning.
3. Understanding and organizing subject matter for student learning.
4. Planning instruction and designing learning experiences for all students.
5. Assessing students for learning.
6. Developing as a professional educator.

The NBPTS framework is based on five core propositions (National Board for Professional Teaching Standards, 2014):

1. Teachers are committed to students and their learning.
2. Teachers know the subjects they teach and how to teach those subjects to students.
3. Teachers are responsible for managing and monitoring student learning.
4. Teachers think systematically about their practice and learn from experience.
5. Teachers are members of learning communities.

Case Study District Procedures

As shown in Table 6.1, Laurel implemented a new standards-based teacher evaluation procedure district-wide in 2010, based on the CSTPs. It was the only of the three case study districts to have fully adopted the CSTPs as the basis for its teacher evaluation system. Laurel’s evaluation required two classroom observations, along with teacher-principal conferences before and after each observation, which exceeds the state’s minimum requirement. The second observation was optional for permanent teachers who received across-the-board satisfactory ratings during the first observation.¹⁵ Teachers in Laurel could, when recommended by an administrator, enter a “professional growth cycle” and be evaluated three times in five years, rather than once every two years. This arrangement is permitted, but not required by the Stull Act. As with the other case study districts, Laurel’s evaluation system used a three-level scale to rate teachers. Teachers received one of three ratings on each of the standards assessed, in addition to an overall rating for the year. The categories were: “satisfactory,” “needs improvement,” or “unsatisfactory.”¹⁶

¹⁵ Laurel has not made any policy changes since the existing policy was implemented in 2010.

¹⁶ These terms were purposely standardized across the three districts to preserve anonymity.

Table 6.1 District Teacher Evaluation Policies Prior to 2010

| | Laurel | Juniper | Sumac |
|--------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Last modified | 2008 | 2010 | 2006 (est.) |
| Implemented District-wide | 2010 | 2009 | 2001 (est.) |
| Evaluation Frequency | <u>Probationary/temporary:</u> once a year <u>Permanent:</u> once every two years, unless rated “unsatisfactory,” then evaluated again the next year <u>Professional growth cycle:</u> 3 times in 5 years, must be recommended | <u>Probationary/temporary:</u> once a year <u>Permanent:</u> once every two years, unless receive “unsatisfactory,” then evaluated each year until improves <u>Permanent teachers after 10 years:</u> if have ratings of “satisfactory,” can be evaluated every 5 years | <u>Probationary/temporary:</u> once a year <u>Permanent:</u> once every two years, unless receive “unsatisfactory” or “needs improvement” rating in any area, then evaluated again the next year |
| Formal Classroom Observations | Two (second is optional for permanent teachers if rated satisfactory in all categories) Must be at least 20 minutes | At least one (additional can be requested) | <u>Probationary/temporary:</u> at least three <u>Permanent:</u> at least one <u>Permanent on remediation plan:</u> if not rated satisfactory, at least three No longer than 60 minutes Only one must be scheduled in advance |
| Conferences | Pre-observation and post-observation (for each observation) Final evaluation conference in the spring | Pre-observation in the fall and final evaluation conference in the spring | Post-observation conference upon request Final evaluation conference (Additional conferences if on remediation plan) |
| Standards | Based on CSTPs Not evaluated on all standards at once (only three per observation) | Not standards-based | CSTPs incorporated loosely |
| Overall Ratings | 3 levels: Satisfactory, Needs Improvement, Unsatisfactory | 3 levels: Satisfactory, Needs Improvement, Unsatisfactory | 3 levels: Satisfactory, Needs Improvement, Unsatisfactory |
| Other Features | Teachers may bring artifacts Annual goal-setting forms required | Mid-year progress report for probationary teachers | |
| Evaluators | Principal or administrator appointed by the principal | Principal or administrator appointed by the principal | Principal or administrator appointed by the principal |
| Alternatives | N/A | Optional alternative to be defined and agreed upon by evaluator and teacher | N/A |
| Consequences | Intervention cycle and remediation plan PAR Recommended for dismissal | Remediation plan More frequent evaluation | Remediation plan More frequent observations Recommended for dismissal (Used to have a PAR process) |

Note: Some of the terms used in this table and related discussion were standardized across districts (where possible) to maintain anonymity and may not match the actual wording in district collective bargaining agreements.

The pre-existing evaluation system in Juniper also employed a three-level rating scale. However, the evaluation forms included only the four categories enumerated in state law and did

not incorporate any of the NBPTS or CSTPs. Juniper’s evaluation process included a pre-conference, one formal classroom observation, and a final conference in the spring. Permanent teachers who have received satisfactory ratings for ten years may be evaluated less frequently – once every five years. The CBA in Juniper allowed for an alternative evaluation process, though it was not well-defined and only used by a few teachers. As part of that process, the teacher would solicit feedback from students and parents, conduct a self-evaluation, and provide a summary to the administrator. Many stakeholders had little knowledge of this option.

As in the other two districts, teachers in Sumac were evaluated on a three-level rating scale. The pre-existing teacher evaluation policy required at least three observations for probationary and temporary teachers, as well as permanent teachers on remediation plans, who had yet to reach the satisfactory level. However, the post-observation conference was offered only upon request. Sumac used to have a PAR program to support struggling teachers, but it was eliminated years ago due to budget cuts. Teachers also received support from BTSA¹⁷ providers and Teachers on Special Assignment (TOSAs). Should teachers not make the targeted improvements, they could be recommended for dismissal. The pre-existing evaluation forms used by Sumac included the four categories required by Stull Act, as well as some additional criteria from the CSTPs.

Common Concerns

Stakeholders voiced several concerns about their pre-existing teacher evaluation policies, some of which included:

- No rating category that goes beyond “satisfactory;”
- Most teachers received the highest rating;
- Low levels of uniformity in implementation; and
- Rubrics were not standards-based.

In general, most stakeholders agreed that their existing systems were not effective tools for improving teacher practice. As one stakeholder said:

Well, I think, our district is no different than any other district in that, principals are overworked and evaluation is very rote... It's more like, "Let's just do it as a formality." But, is it really leading to teacher development? No.

Across the board, stakeholders felt that having a system that helps all teachers grow professionally was important. They also agreed that there were a small number of low-performing teachers that needed extra support.

Interviewees noted that many teachers were unhappy that the existing evaluation systems did not recognize excellence; the highest rating teachers could receive was “satisfactory,” confirming that they met standards but not distinguishing those who exceeded the standards. Given that nearly all teachers were receiving the rating of “satisfactory,” the existing evaluations failed to

¹⁷ BTSA is California’s new teacher induction program.

differentiate among teachers. Stakeholders also noted dissatisfaction regarding how classroom observations were conducted. Some individuals described classroom observations as a “dog and pony show,” saying that they only provided a snapshot and did not accurately represent a teacher’s day-to-day performance. Some stakeholders felt that principals were not in classrooms for enough time to make an accurate assessment of teacher performance. As one individual in Sumac said:

The administrator might be in your room, 10 minutes, or maybe 15 minutes, and then they leave and then months later, you would get the write up of what they saw.

Another stakeholder criticized the system for a lack of teacher-administrator dialogue:

So there was no professional dialogue. There was no anything. There was no goal-setting. There was no, "What are you trying to achieve by this lesson? What are the long-term goals here?" There was none of that thing that was happening across the board. It was just... It happened in little tiny pockets, at little tiny schools, in little tiny settings. And that was really about it.

Teachers wanted the process to be more reflective. A stakeholder in Sumac pointed out that the evaluation did not require or assess lesson-planning, which was noted as a weakness.

Additionally, stakeholders in all three districts expressed concern about the consistency of implementation. One individual in Juniper described specific problems with their pre-existing evaluation form:

It has areas for comments and whether you are [satisfactory], needs improvement, or unsatisfactory. And so, it has nothing to do with the teacher-initiated standards that are out there, that we think are so good. It meant that employees get wildly different feedback depending on their evaluator. And, it was not really a professional growth tool that we could call a consistent tool.

Furthermore, some teachers perceived the evaluation to be a “gotcha,” assuming that the true purpose was to catch them doing something poorly and penalize them. Administrators were aware of this perception and hoped to clarify that their goal was not to be punitive.

In spite of these concerns, not all stakeholders believed that change was necessary. In higher-performing districts, it was more difficult to make the argument that reform was urgent or essential. As one individual said:

Is evaluation needed? Yes, politically. Is it really needed? Do we really need to change it for this district if we're top performing in the state?

In lower-performing districts, like Sumac, there was more pressure to show the community that the district was taking action to improve teacher performance and student achievement.

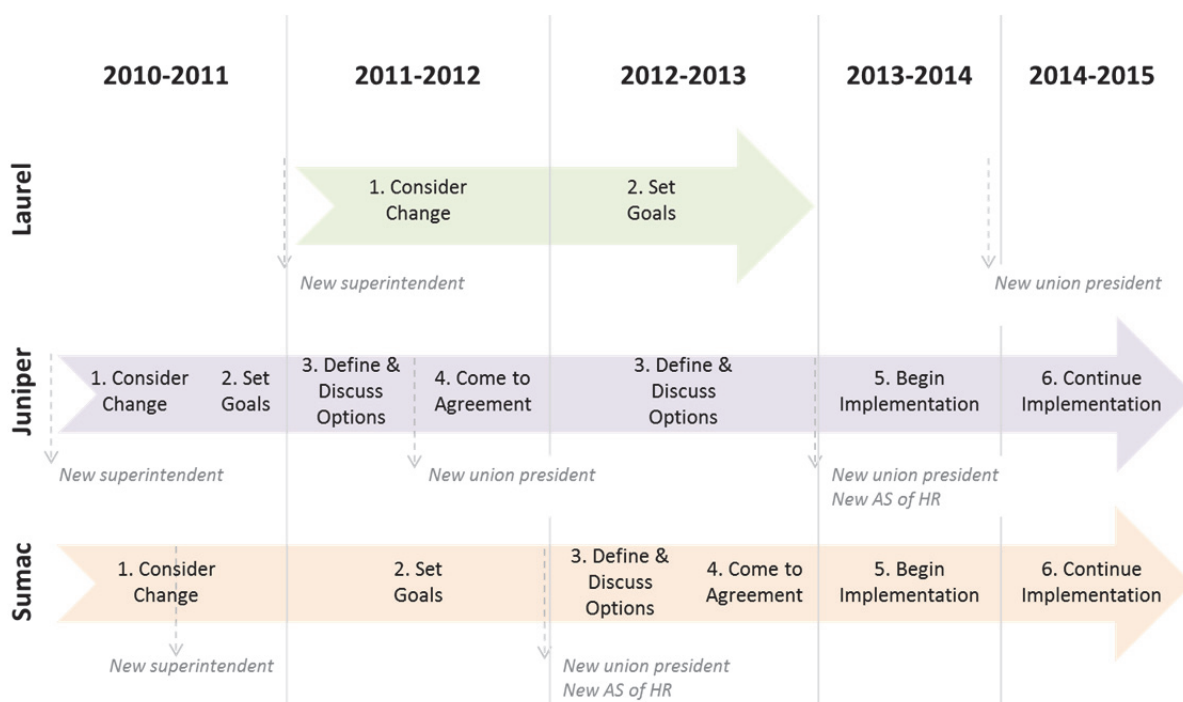
The evaluation systems in the three case study districts (under both the pre-existing and newly adopted policies) did not meaningfully utilize student achievement data; a teacher’s overall rating was not affected by measures of student achievement. While student progress toward achievement objectives might be discussed as part of the evaluation process, teachers

were not judged directly on student performance. However, the majority of district stakeholders did not identify this as a weakness. A few administrators did say that they would like to see their districts make better use of student achievement data, in general.

District Overviews

As shown in Figure 6.1, the three case study districts began their reform efforts around the same time, though they are currently at different stages. A common start time is not surprising given that high profile reforms to teacher evaluation policies were spreading nationwide and drawing stakeholder attention. All three districts hired new superintendents in either 2010 or 2011, when current efforts to revise teacher evaluation policies began. They also experienced turnover in union and other district leadership throughout the reform process, which is denoted by the arrows in the figure below.

Figure 6.1 District Reform Timeline



While districts began considering change around the same time and followed a similar reform process, their stories and progress are unique. The character of labor-management relationships in each district shaped their progress and policies. As noted in the literature, the success of joint district-union reform efforts depends on the ability of stakeholders to work together effectively. Each of the case study districts had experienced labor-management conflict in the past, some of which still influences labor-management relations today. When reform efforts began, some stakeholders in one district noted that the institutional memory of one strike

that occurred decades before had persisted. In another district, there was some residual bitterness among secondary teachers who were not satisfied with the concessions they had to make when the district unified decades ago. The entire community later lost a degree of trust in the district when a long-tenured superintendent had to retire in the midst of a scandal. Labor-management relationships were not been stagnant and all three districts experienced ups and downs over the years. I provide a summary of each district's reform efforts and labor-management relationships below.

Laurel

Laurel's case is best described as a collaborative effort that lost momentum. Laurel previously adopted a standards-based teacher evaluation system that was implemented district-wide in the 2010-2011 school year, after a lengthy reform process that began in 2004. However, there were many policy changes that district stakeholders still wanted to see. Current efforts to change the district's teacher evaluation system began in 2011. The new superintendent made labor-management collaboration a district-wide goal and began a transition to interest-based bargaining (IBB). Reform efforts were jointly initiated by district and union leadership, and the first stage, consider change, spanned an entire year. The district went on to form a joint labor-management teacher evaluation committee in 2012,¹⁸ which met monthly over the course of the 2012-2013 school year. After a year in the second goal-setting stage, the committee produced a list of recommendations for changing their existing teacher evaluation policy. However, as of December, 2014 the district had not made any progress; the committee's final meeting was in May, 2013. Stakeholders considered the work to be "exploratory" and even "leisurely," with no concrete timeline or deadlines in place. While district and union stakeholders worked well together, the lack of collective urgency stalled the process. At the conclusion of this research, Laurel had yet to transition into the next stage of defining concrete policy options.

Juniper

Juniper's case is one of progress, punctuated by setbacks. Of the three case study districts, Juniper struggled the most with labor-management relations. Juniper was faced with a pre-existing climate of distrust that was difficult to overcome and substantial turnover in both district and union leadership. The current superintendent joined the district in 2010 and three different individuals served as president of the teachers union from 2010-2015. Historically, it has been a challenge to find individuals to run for union leadership positions. The district also hired a new Assistant Superintendent of HR in 2013, while their reform efforts were already underway. Still, there is some precedent for district-union collaboration in Juniper, as well. Prior to 2010, the previous superintendent and the union president at the time (UP #1) had a good working relationship. Interviewees noted that they communicated well and organized many joint district-

¹⁸ This was not the same committee that was formed in 2004 and designed the existing teacher evaluation policy.

union presentations and committees. The relationship between Juniper's superintendent and union leadership has been different under each of the union's presidents. Stakeholders reported that, as of the second year of the pilot, labor-management relations were improving. However, the relationship between the Juniper's school board and teachers union was not very positive. The union had several objections to how the Board's work was conducted and some viewed the school board as anti-union and anti-tenure, which made teachers concerned about their job security.

Compared to the other districts, Juniper moved more quickly through the first four stages, forming its first teacher evaluation committee in the fall of 2010 and coming to a tentative agreement by the end of the 2011-2012 school year. However, the district then experienced some setbacks and was not able to begin a two-year pilot until 2013. When new union leadership took office in 2012, the membership of the teacher evaluation committee changed and labor-management relations became strained. That second committee backtracked and spent another year in the "define and discuss options" stage, before agreeing to move forward with a teacher evaluation pilot. The pilot was optional for permanent teachers and support for the new system varied by school site. District administrators solicited stakeholder feedback during the first year of the pilot, though Juniper's second teacher evaluation committee did not continue to meet that year and no policy changes were made. Despite some implementation challenges, and a contentious round of contract negotiations, Juniper continued with implementation in 2014-2015 and will form a third committee to evaluate the pilot and propose recommendations. The third committee will meet throughout the second year of implementation and then make recommendations to the school board in the spring of 2015. Any proposed changes must be approved by five out of eight committee members before going to negotiations. Some anticipated that those negotiations would be difficult:

I think it will be a make or break issue for our Board next year. It'll be a huge issue for them. So, it could really drag negotiations on. I hope we can do it in a positive way, but it'll be difficult.

The committee will look to improve the goal-setting process built-in to the evaluation system, to make it more targeted. As one stakeholder described:

So in our first meeting, we were really looking at the goal development piece and how to get teachers, 'cause within each standard there are, I don't know if you remember the forms. There's a number of elements with a standard and so what we're talking about is how do we get teachers to be specific about the elements that they select? And so we just changed a little bit of the wording so that when they do select a standard, they can highlight some of their work and the elements more specifically.

Both the district and union representatives on the committee agreed to those changes. They also want to work with principals on coaching and PD, to make the evaluation process an effective tool for professional growth:

I mean, the main thing is that we've been working a lot with principals and talking a lot with principals about being a coaching principal. So that evaluation just doesn't... I mean this idea of having discussions around instruction don't just consist within that container of evaluation, it's ongoing. The principals are regularly in classrooms, regularly engaging in these types of conversations, and their goal is to support teachers in moving the work forward. And so I think that's one of the big things right now.

As of 2014, there was still disagreement over the language used in some of the standards that will need to be resolved as implementation continues.

Sumac

Sumac is a case of the “perfect storm” for a collaborative agreement. Sumac has historically had a strong teachers union that held significant political clout. In the past, the labor-management interactions were often combative. Several stakeholders described district culture as a “them v. us” environment, and noted that both the district and the union were often referred to as “old boys’ clubs.” Some described Sumac’s previous Assistant Superintendent of HR as being from the “old style” of labor-management relations. As one stakeholder in Sumac noted:

The relationship used to be very strained and it used to be very... I wouldn't say combative, but I would say... It just wasn't, it wasn't really harmonious. It wasn't a lot of collaboration. There wasn't a lot of working together.

However, in 2011, with the arrival of new district leadership, the labor-management relationship began and continued to improve.

Sumac switched to IBB several decades prior to the beginning of their teacher evaluation reform efforts in 2010. The superintendent at the time made the decision in response to the numerous fights they were having at the bargaining table. Though it was not initially a mutual decision, both parties became committed to it over time. Though IBB in and of itself does not repair relationships or eliminate conflict, stakeholders in Sumac said there have been some positive effects. In fact, stakeholders noted that they try to apply the IBB approach of focusing on issues and interests, rather than positions, outside of bargaining.

Sumac experienced what stakeholders referred to as a “perfect storm” for a collaborative reform effort. The district formally established its teacher evaluation committee in the fall of 2011, a year later than Juniper did. However, they were still able to complete the first year of a teacher evaluation pilot in 2013-2014. The new superintendent, who arrived while the district was already considering change to its teacher evaluation policy, made labor-management collaboration a priority. Union leadership was interested in reform and proactive in establishing Sumac’s teacher evaluation committee. The committee invested significant time – a full year – in the goal-setting stage. The group agreed to a set of guiding principles and developed common language. Stakeholders described the process as being difficult, as the group worked to build relationships and establish trust. However, that investment allowed them to move more quickly through the following two stages, and the district and union agreed to a pilot evaluation program

without a formal MOU. The committee strategically selected five pilot schools that had positive teacher-administrator relationships, and began implementation in 2013.

The pilot continued throughout the 2014-15 school year with participants from all school sites. The committee engaged the district community throughout the reform process, issuing joint-communications and soliciting stakeholder feedback. In response, the committee made some modifications to the policy prior to the second year of implementation. The district addressed stakeholder concerns about the time commitment for administrators, noting that they intended to add additional Assistant Principal positions in the second year of implementation. They also decided to combine some of the teacher-administrator conferences and shorten the observations required for exemplary teachers. For example, the pilot prescribed separate goal-setting and pre-observation meetings to be scheduled at the beginning of the year, but participants found that those could be done at the same time. While they wanted to be mindful of administrator time, they were wary of curtailing the process to a point where it became a “rubber stamp” and lost its value. Sumac is also considering adding a five-year skip for veteran teachers with positive evaluations and an alternative evaluation that would be agreed upon by a teacher and principal.

As their work progressed, labor-management relations improved, and the committee served as a model of labor-management collaboration for the entire district. Once the new evaluation policy became a permanent fixture in the CBA, at the end of the first year of implementation, they began referring to it as a “structured roll-out,” rather than a pilot. Stakeholders in Sumac planned to monitor consistency in implementation across school sites and have the committee meet less often. The committee planned to make improvements to the system before it is mandatory for all teachers in 2015-2016. Any changes proposed by the committee will have to be approved in contract negotiations.

Organization of Results

In the following four chapters – seven, eight, nine, and ten – I present my detailed findings for the case study districts, organized by the stages in the reform process, as outlined in my conceptual framework. (My conclusions and recommendations follow in Chapter 11.) For each reform stage, I compare the actions of the three districts, highlighting notable similarities and differences and derive answers to my three research questions. The first chapter, “Setting the Stage” provides an overview of the motivation for district reform efforts, as well as the first stage in the reform process: “consider change.” The second chapter, “Approach to Reform,” addresses the next reform stage, “set goals,” and describes the formation of district teacher evaluation committees. The third chapter, “Hammering Out the Details,” contains two stages: “define and discuss options” and “come to agreement.” Finally, the fourth chapter, “Moving Forward,” includes the results for the two implementation stages and details the evaluation pilots in Juniper

and Sumac. Laurel is only included in the first two results chapters, as the district did not move beyond the goal-setting stage.

Each results chapter is organized into three sections that match my three research questions:

- Stakeholder roles;
- Stakeholder relationships; and
- Influential factors.

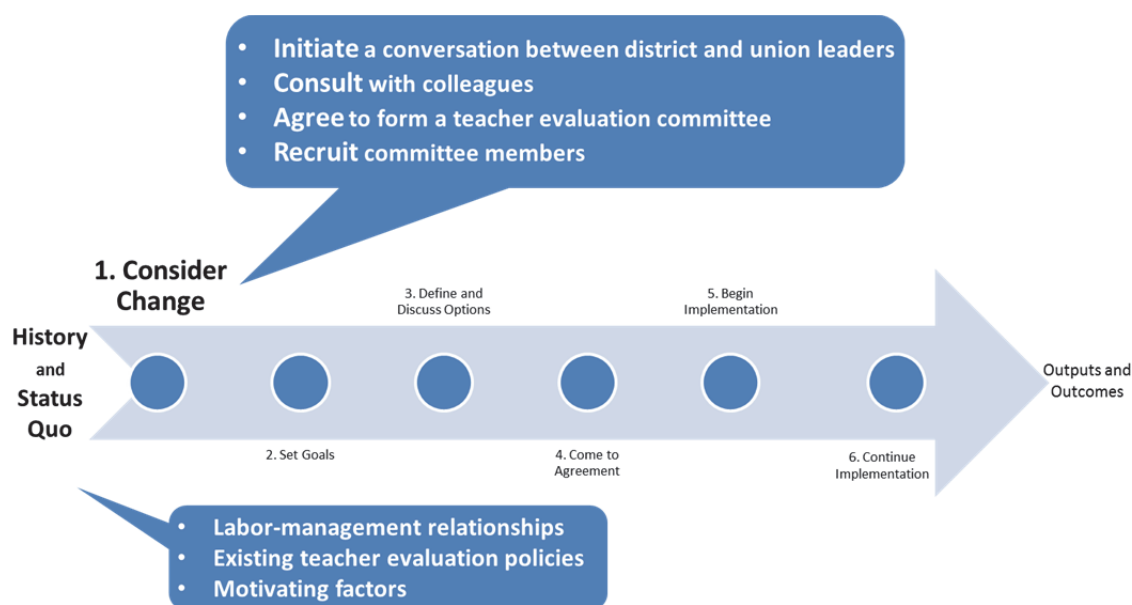
I first summarize the reform activities in the three districts during the related stages. From those descriptions, I identify the stakeholder roles and relationships, which align with my first two research questions. I then address the factors that influence those roles and relationships, as well as what the districts were able to accomplish during each stage. This discussion informs my third research question. Following my presentation of results, I conclude my dissertation by addressing each of the three research questions in Chapter 11. I discuss the most important lessons learned and make policy recommendations.

Chapter Seven: Setting the Stage

In this first results section, I set the stage for how each district’s teacher evaluation reform efforts began. I review the key activities that defined the first reform stage, “consider change,” and the factors that motivated this work. I summarize the roles stakeholders played in considering change and deciding to move forward with the current reform efforts, and then describe that character of the labor-management relations observed. Lastly, I employ cross-case analysis to identify the most influential factors that shaped the beginning of the reform process.

Figure 7.1 outlines the key reform activities during this stage.

Figure 7.1 Reform Activities – Setting the Stage



Though not a stage, each district’s story began with the history and status quo of their labor-management relationships and teacher evaluation policies, which were discussed in the previous chapter. In the first reform stage, leaders in all three districts initiated conversations about their concerns. Sometimes these started during contract negotiations; other conversations were more informal. After consulting with their colleagues, district and union leaders agreed to form joint labor-management committees that would spearhead their teacher evaluation reform efforts. Superintendents appointed district representatives to these committees, while union leaders recruited teacher representatives to participate.

Districts were motivated to pursue this work for similar reasons. Common motivating factors included:

- A general dissatisfaction with the status quo, including existing evaluation policies;
- The desire to be proactive in a changing policy environment; and
- New leadership.

Perhaps the strongest motivator was an internal dissatisfaction with the status quo of each district's teacher evaluation system, student performance, or even the labor-management relationship, in general. One stakeholder in Sumac gave a concise summary of why it was the right time for her district to consider change:

One, [the union] got a new president... Several, if not most districts around or in [California] were updating or have been updating theirs. The form used in this district is very outdated and I think everyone pretty much agrees with that. So it was time and under the new superintendent and the new president I think, it just kind of planets aligned and it was the right time.

In other words, it was the “perfect storm,” as more than one of her colleagues articulated. Similar themes emerged in all three districts.

In addition to having concerns about their existing policies, stakeholders were aware that the national policy landscape was changing. Media coverage of teacher evaluation reform in the mid-2000s sparked conversations across the education community and in all three of the case study districts. Debates over policies related to teacher evaluation, especially about student growth measures, merit pay, and teacher tenure, also drew attention and concern.

Though California was not mandating changes to district teacher evaluation policies in 2010, when the case study districts began considering change, the threat of upcoming legislation was ever-present. In addition, many stakeholders were aware that they might not be eligible for future funding opportunities unless they changed their evaluation policies. The combination of these factors fueled a desire to be proactive, especially among union leaders, who did not want to wait for policies to be imposed on them. As one individual said:

I think the idea... of being more proactive and just not waiting and not waiting 'til something bad happens, and then being told, "Oh, okay, now because this is horrible, this is what you're going to have to do." I mean, given the chance that we can collaborate to craft something new and better, I think is really highly engaging and motivating.

District stakeholders wanted to have a hand in shaping the path their districts would eventually take on teacher evaluation, unlike what they observed happening in other districts. However, not all stakeholders were equally concerned about moving forward and did not see changes to state legislation as imminent. While some may have preferred to be proactive in initiating local reform, others seemed more comfortable with the status quo, as one individual said:

You got some reform-minded people within the movement who just think, "We got to get ahead of this", or some others who say, "Look, we just got to hold the line. The more we allow this conversation to move forward, the more we're going

to be losing control." There's a tactical conversation within labor and even within our union, in our school district, I've seen different approaches to this question.

A small number of individuals interviewed said that media coverage of teacher evaluation and related reforms was not a significant driving factor and that the real source of motivation was an internal desire to improve their own evaluation systems and better support teachers.

Many stakeholders identified the arrival of new superintendents as factor that motivated their districts' teacher evaluation reform efforts. These new district leaders had worked on teacher evaluations in their previous districts and identified it as an area of growth in the case study districts. Had the superintendents not made this work a priority, it may have been more difficult to mobilize stakeholders to consider change. Some individuals perceived their superintendent's interest in teacher evaluation reform, in conjunction with other initiatives, as an attempt to have a measurable impact in their new position. Superintendents were also responding to the expectations of school board members, who set over-arching goals for the district. For the most part, superintendents were not receiving pressure from the community to change teacher evaluation policy. However, stakeholders in Sumac noted that their new superintendent did have to respond to general external criticism of the district's performance.

Stakeholder Roles

There were several different stakeholders who participated in this early stage of "considering change." The most involved stakeholders were the district and union leaders. Other district administrators, union representatives, and sometimes school administrators and school board members participated at this stage, as well. The specific stakeholder roles most commonly observed as these three districts began considering change include:

- *Initiators* who got the conversation about teacher evaluation reform started;
- *Conveners* who brought a group of stakeholders together;
- *Contributors* who agreed to join the district's reform efforts; and
- *Supporters* who paid attention and expressed interest, but were not directly involved.¹⁹

Tables 7.1 and 7.2 summarize the stakeholders involved in the first stage of the reform process and the roles they played in each district, with examples of how they contributed to the work. (All stakeholder roles were defined earlier in Chapter 4.)

¹⁹ Though some "supporters" are identified, this does not mean that there were no stakeholders who did not support this work. There were likely "objectors" but they may not have been vocal this early on in the process.

Table 7.1 District Stakeholder Roles in Considering Change

| | Laurel Unified | Juniper Unified | Sumac Unified |
|---------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Superintendent | <p>Initiator Set a goal of collaboration at beginning of year and had conversations with UP.</p> <p>Convener Helped convene first ad hoc meeting and scheduled the first meeting of the full committee. Recruited district administrators.</p> | <p>Initiator Approached the union about forming a committee. Organized trip to a peer district and proposed their system as a model.</p> | <p>Initiators <i>Previous super</i> helped get the initial conversation started. <i>Current super</i> communicated this would be a district priority.</p> <p>Conveners <i>Previous super</i> helped convene the first exploratory committee. <i>Current super</i> brought a larger committee together and appointed district admin.</p> |
| Assistant Superintendent of Human Resources | <p>Convener Involved in conversations with UP about how to move forward. Recruited members with HR and helped convene the committee. Served as a liaison between superintendent and school board. Attended ad hoc meeting.</p> | | <p>Contributor <i>Previous AS of HR</i> was interested in the topic and was consulted on committee membership.</p> |
| Other district admin | <p>Contributors Some attended initial ad hoc meeting.</p> | <p>Contributors Some involved in the negotiations that led to the formation of the committee.</p> | <p>Contributor Was consulted on committee membership.</p> |
| School admin | | | <p>Supporters One school independently piloted a new evaluation system because they received a grant.</p> |
| School Board | <p>Contributors One member was very interested and helped organize ad hoc meeting. Two members were involved in ad-hoc meeting and talked to the super about moving forward.</p> <p>Supporters Other members also wanted to be proactive on teacher evaluations.</p> | <p>Supporters Members communicated to the new super that evaluation reform was a goal, but were not directly involved in the work. One board member and a parent offered to be on the committee, but it was decided that it was better if left to teachers and administrators.</p> | |

Table 7.2 Teachers Union Stakeholders Roles in Considering Change

| | Laurel Unified | Juniper Unified | Sumac Unified |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Union President | <p>Initiator Initiated conversation with union leadership, then district admin. After the ad hoc meeting, asked super about starting committee.</p> <p>Convener Helped convene ad hoc meeting.</p> | <p>Convener <i>Union President #1</i> agreed with the superintendent to form a committee. Recruited union reps to sit on the committee.</p> | <p>Contributor <i>Previous union president</i> agreed to move forward. Was consulted on committee membership and chose another union rep to be a committee co-chair.²⁰</p> |
| Union reps/ teachers | <p>Initiator One rep also approached superintendent about getting committee started.</p> <p>Contributors Union reps met and chose this as a priority. Attended ad hoc meeting and involved in early conversations. Helped recruit teachers to be on committee.</p> | | |
| Union reps / teachers | <p>Contributors Union reps met and chose this as a priority. Attended ad hoc meeting and involved in early conversations. Helped recruit teachers to be on committee.</p> <p>Contributors Some attended the initial ad hoc meeting.</p> | <p>Supporters Some involved in early conversations, but did not play an active role until the committee was formed.</p> | <p>Initiator The union negotiator suggested forming a committee.</p> <p>Contributors Union reps were consulted on committee membership.</p> |

Initiators

In general, there was no one person who was wholly responsible for initiating any district's reform efforts. Superintendents and union leaders, including presidents and some representatives, worked together to initiate change. The idea to pursue reform was not a district mandate; union leaders were proactive early in the process, especially in Laurel and Sumac. In Laurel, the superintendent, union president, and one particular union representative served as the “initiators” of their reform efforts. All three stakeholder groups – union, district, and school board – shared an interest in teacher evaluation reform. As one stakeholder described:

There was a real interest on the part of our board of education in particular. I think the union had interest as well. I think it kind of came from all three prongs, really, the union, management, and the board of education, to examine whether what we were doing is effective or ineffective, whether our process worked.

After the idea of convening a formal committee to research teacher evaluation policies emerged in an ad hoc meeting, union leaders met over the summer and decided that they wanted to be proactive in these reform efforts. As one union representative in Laurel said:

We really wanted to be able to be a little thoughtful and more proactive than reactive. That was kind of our mantra, that we wanted to really lead rather than follow. We wanted to really do some thoughtful work instead of waiting maybe for something to come down from administration to us. We kind of wanted to

lead this, so we asked our, and she was brand new, our new superintendent, what she thought about that and she thought that would be great.

The union president and one interested representative approached the superintendent about getting the formal committee started, which the superintendent agreed to do.

In Sumac, the previous superintendent brought up the idea of reviewing the teacher evaluation documents at a monthly meeting with the teachers association, and that is where the conversation began. A union representative, who was also a member of the negotiations team, suggested that they form a committee. The decision to move forward was mutual. When the current superintendent arrived, she communicated that this work would be a district priority and kept the efforts moving forward. In Juniper, the newly-hired superintendent was the key initiator. She approached the union president about organizing a committee and the union president agreed to help lead it.

Conveners

The individuals who were involved in convening their districts' teacher evaluation committees were thoughtful about who the members should be. In general, superintendents appointed the district representatives and union leaders recruited the teacher representatives. In Laurel, the Assistant Superintendent of HR took on a leadership role at this point and helped to recruit participants and organize the first committee meeting. The teachers unions in Laurel and Juniper sent an open invitation to all teachers to join the committee. Sumac recruited committee members from teachers who already held elected positions in the union. While union leaders had to rely on teacher volunteers, they emphasized that it was important to have representation across all grade levels – elementary, middle, and high school. Likewise, it was important to include district administrators with different areas of expertise.

Contributors and Supporters

As shown in Tables 7.1 and 7.2, several other stakeholders contributed to this early stage in the reform process. Many district administrators and union representatives participated in early conversations and the decision to consider change, either by attending informal meetings or agreeing to be involved once committees were formed. School board members and school administrators were marginally, if at all, involved. Some school board members did encourage reform efforts, though they mostly acted as observers. They identified teacher evaluation reform as a district goal, but were not ultimately involved in the teacher evaluation committees. One school board member in Laurel was responsible for bringing in the outside consultant, who attended the first ad hoc meeting; the school board member attended the meeting as well. However, after that meeting, he took a backseat. A few district stakeholders, who would later play significant roles in the reform process, were not yet at their districts when these efforts began. These included the current Assistant Superintendents of HR in both Juniper and Sumac.

Stakeholder Relationships

As discussed earlier, labor-management relationships have varied over time, maintain a deep connection to district history, and are affected by changes in leadership. The three relationship categories I adopted in my conceptual framework are collaborative, civil, and contentious. (See Table 4.5) At the beginning of the first stage in the reform process, each of the three case study districts exhibited different types of relationships, yet all were able to come to an agreement to form a teacher evaluation committee. Laurel started with the most positive relationship, whereas stakeholders in Juniper and Sumac had initial issues with trust. While considering change, Laurel exhibited the most collaborative interactions, whereas the labor-management interactions in Juniper and Sumac were best described as “civil.”

When the current superintendent of Laurel joined the district in 2011, she placed a priority on improving labor-management relations and deliberately set the tone for a culture of collaboration. One individual noted that there was an overall “collaborative vibe” in the district at the time. To that end, she led the district in transitioning from traditional contract negotiations to IBB. Because it was still a newly-adopted process, the IBB approach was not yet as engrained into the district culture in Laurel as it was in Sumac. Laurel’s teachers union had been very active and, overall, stakeholders described the labor-management relationship as good. Laurel’s union leaders and superintendent met regularly. Stakeholders saw that as a unique, and positive, aspect of their district’s culture. The district and union had worked together on several other initiatives, including their LCAP²⁰ and the implementation of the new Common Core standards. Some stakeholders described a widespread “committee culture” in Laurel Unified. The district and union leadership had even worked well together on difficult issues, such as dismissing teachers. One stakeholder said:

We tend to be a district that’s more focused on consensus, on really working together. It’s not that we don’t disagree on some things; we do.

As a result, the president of the teachers union and the superintendent developed a positive working relationship, which enabled them to have productive conversations about changing their teacher evaluation policy. The decision to form a committee was perceived as a joint initiative, not something being imposed by the district or demanded by the union. In fact, several stakeholders interviewed could not recall which party first proposed the idea.

However, not all stakeholders viewed the relationship as entirely positive, and some leaders in Laurel were more willing to collaborate than others. Some barriers to a productive labor-

²⁰ In 2013, California changed the way it funds schools, implementing the new Local Control Funding Formula (LCFF). As part of the new funding system, each district had to create Local Control and Accountability Plan (LCAP), in which they would develop annual goals for students that align with state priorities and detail how they planned to meet those goals. The LCAP would address a three-year period and be updated annually. Teachers, administrators, parents, students, and all bargaining units must be consulted when developing the plan. Districts were required to submit the first LCAP for review beginning July 1, 2014, which coincided with the case study districts’ work on teacher evaluation (California Department of Education, 2015).

management relationship remained. One stakeholder observed that some union members retained the mentality that “collaborating with the district is selling out,” and that may put pressure on union leadership to employ more traditional bargaining tactics. Still, most stakeholders interviewed believed that the relationship they had was better than what they have observed in many other districts, where adversarial interactions were the norm.

While interactions in Juniper and Sumac were not collaborative when the teacher evaluation reform efforts began, district and union leaders did not act as adversaries and were able to cooperate to the degree necessary to move forward to the next stage of reform. In Juniper, there seemed to be an environment of limited trust. The new superintendent made an effort to form many joint labor-management committees, perhaps as a strategy to restore that trust. As one individual said:

I think that just kind of all came together and he wanted to start a committee and so we have just a boatload of committees those years because it was all about working together and trying to find commonalities so that we can create positive change.

While the district and union agreed to explore change, some stakeholders remained skeptical of the superintendent’s motivation and were reluctant to proceed.

Similarly, stakeholders in Sumac successfully negotiated the formation of a teacher evaluation committee, even though some union members remained reluctant. The relationship improved when new leaders took office, including the superintendent, her cabinet, the union president, and some school board members. Like the superintendent of Laurel, the new superintendent in Sumac was vocal about promoting collaboration, which helped put the labor-management relationship on a positive trajectory. She wanted to be a “team player” and deliberately used the term, “team,” frequently. Aware that the district’s labor-management relations had a contentious history, she made it a priority to improve that relationship. Though Sumac’s teachers union leadership may have been skeptical of the many personnel changes in the district office at first, but would soon form positive working relationships with the new district leaders.

Influential Factors

All three districts were able to move forward from this initial stage of the reform process by agreeing to consider change and forming joint labor-management committees to explore policy options. There were many similarities, but also differences, in their approaches and the circumstances they faced, and the speed and ease with which districts progressed varied. Cross-case comparisons revealed several factors that either enabled or impeded progress at this stage. In this section, I present five key influences that affected labor-management interactions and the ability of each district to move forward from considering change. These factors are related to

leadership, district culture and characteristics, the approach to reform, and the timing of their efforts.

Leaders Provided Momentum for the Work

Changes in district leadership facilitated progress and added momentum. Superintendents and union leaders were instrumental in both initiating district reform efforts and convening teacher evaluation committees. Stakeholders in all districts noted the influence their district's leaders had on the efforts to change teacher evaluation policies, as well as on the character of labor-management interactions at the beginning of the reform process. Further, the willingness of these leaders to work together was a key factor that enabled districts to meaningfully consider change. It is possible that these efforts would not have gotten off the ground under different leadership. As one individual in Laurel said:

I have to say honestly, even though it shouldn't really, in my opinion, depend on the person, it sometimes does. But this particular superintendent really wants to work as a team and so does the association, so that's how we try to solve things.

Similarly, one stakeholder in Sumac reflected on the importance of the union president, who had previous experience with developing teacher evaluation policies and a long tenure in the district:

Again, I would just go back to the president of [the teachers union] was new, our superintendent was new... I don't know if under the former president it would have gone as smoothly. And certainly under former superintendents I don't know that it would have gone as smoothly.

Not surprisingly, these leaders had the largest influence on the initiation of district reform efforts. Leadership turnover can create challenges, but also provides a natural opportunity for considering change. It is clear that the new superintendents in all three districts championed the work around teacher evaluation policies. Some stakeholders believed that their new superintendents were looking to make their mark on the district, which is why they were interested in exploring new policies.

Shared interest in teacher evaluations

Given that California was not mandating changes to district teacher evaluation policies, it was up to district leaders to challenge the status quo. From my interviews with district stakeholders, it was clear that superintendents, union leaders, and in some cases, school board members believed that reforming teacher evaluations had the potential to improve teaching and learning in their districts. Further, these individuals had a desire to be proactive. This was a crucial factor, given that districts have a number of goals and competing priorities. At the time, districts were tasked with a number of simultaneous policy mandates, including the implementation of the Common Core standards and drafting their LCAPs. While there were many individuals who had become interested in teacher evaluation policies, ultimately, it was up to these leaders to make reforming teacher evaluations a policy priority.

Pre-existing relationships

A good working relationship between the superintendent and union president facilitated the early work around teacher evaluations. In all three case study districts, labor-management relationships were at least civil when reform efforts began, and I suspect that districts would not have agreed to move forward if the status quo had been defined by contentious and adversarial interactions. However, the character of those relationships is always evolving, and would not remain the same over the course of the reform process. Positive relationships can be difficult to maintain and may deteriorate or strengthen over time. Current district and union leaders made a concerted effort to foster positive, or at least functional, working relationships.

Communication

In all three districts, leaders from the district and union made an effort to communicate with each other, even scheduling regular meetings. They were willing to meet and talk outside of formal contract negotiations. As one individual said of union leaders:

So I think it's a testament to the desire to do something proactively on the part of our union that they were willing to meet in that format instead.

It was important for labor and management to be on the same page early in the reform process, recognizing that they would have to engage in difficult conversations once the work began. The fact that they did not wait until the next round of contract negotiations to initiate conversations about teacher evaluation, may have set them up for success. By maintaining a dialogue, they were able to agree on priorities. As one district leader said:

So during that time I worked really hard on various different levels to make sure that we were kind of working together toward the same goal, which creates an open honest conversation.

While disagreements still arose, having open conversations helped leaders understand that their counterparts had good intentions. Further, by working together early on, the process was not perceived as a one-sided mandate, which laid the foundation for future collaboration.

Persistence

Though teacher evaluation reform was a priority for many leaders, not all stakeholders agreed with their decisions to pursue change. Leaders must respond to criticism and persist even when met with resistance, which can be especially challenging for union leaders who are elected representatives. Union leaders in all three districts were willing to be proactive and work with management, in spite of the reservations of some of their constituents. In Laurel, the teacher evaluation committee may not have re-started without the insistence of union leaders. Thus, the persistence of leaders proved to be an important factor.

New Superintendents Set the Tone for Labor-Management Interactions

While superintendents are not the only influential leaders, superintendents set district priorities, appoint other leaders, and are ultimately responsible for district functions. As noted above, all superintendents made teacher evaluations a priority which helped the work move forward. Many stakeholders commented on the language their superintendents used, even quoting specific catchphrases or slogans, which were often related to collaboration and excellence. Superintendents were very deliberate in communicating their intentions to work collaboratively. As one stakeholder noted:

I would say that we try, when possible, where we all agree on issues to work as a team, and the superintendent even uses the word "team" all the time.

In Sumac, stakeholders commented that the superintendent was careful about who she brought on her team; the new district leaders all showed a willingness to collaborate. While superintendents set goals and priorities, the superintendents of the case study districts did not want to make top-down policy decisions. To that end, they preferred to form joint-labor management committees. Superintendents also invited union leaders to be involved in a number of different initiatives, not only in the work around teacher evaluations.

District Culture Affected Stakeholder Willingness to Consider Change

History and culture

District history and culture influenced labor-management interactions and a district's approach to policy reform. Efforts to consider change to any policy can be hindered by wounds from past conflicts and persistent distrust, or enabled by positive experiences with past initiatives. In Sumac, stakeholders described being tired with a history of conflict and therefore were invested in improving labor-management interactions:

People with long tenure in the district were fed up with how things were in the past. It was also helpful that some leaders had been in the district a long time and seen the history.

Though they had experienced conflict in the past, they seemed to be ready to work together:

And it's just kind of one of those things where, I think, everyone's philosophies have aligned. And that hasn't always been the history here. There have been combative cabinets and combative union leaderships in the past.

As observed in the case study districts, past conflict can present obstacles, but it does not prohibit future collaboration.

Previous work in the area of teacher evaluations may have helped get current efforts started in Laurel. The district already successfully transitioned to a standards-based evaluation system without outside pressure and independent of any external mandates. That experience ignited mutual interest among internal stakeholders, which may have enabled them to move forward in a positive manner. Experience with IBB or participation in IBB training, which focuses on

consensus, also had a positive effect on interactions in two districts. Sumac has been using IBB for many years. Their decision to consider change to the district's teacher evaluation policy came out of the IBB process. Stakeholders noted that it was part of district culture to not engage in confrontational bargaining, and that the IBB process had helped them build a more collaborative relationship. One stakeholder noted how that progressive approach to bargaining improved their ability to work together to solve problems and commented on the approach they took:

Well what's the issue that we have, and how can we work on that together to come to a resolution that benefits everybody, and most importantly, benefits the needs of our kids within our district?

Laurel began transitioning IBB when current reform efforts were initiated, and stakeholders observed how it affected the way labor and management worked together in other areas. One stakeholder recognized that the positive relationship may make them more open to reform:

I think it's getting more, less positional and more of how can we get things resolved together because of the IBB training that we participated in together.

While the new approach helped in some ways, stakeholders in Laurel still experienced some traditional labor-management issues, such as grievances. It is important to consider that adopting IBB does not make a district immune to conflict or disagreement; however, it may help them develop strategies to work through those issues.

District performance

District culture and labor-management interactions were also influenced by student performance. In fact, a history of low student achievement may have made stakeholders more willing to work together on policy changes. Conversely, high student achievement levels served as a barrier. In a lower-performing district, there may be a shared belief that the status quo is sub-par, while in a higher-performing district, it may be harder to justify a need for change. These patterns were evident when comparing stakeholder reactions to reform efforts in the case study districts.

Sumac has historically been a low-performing school district. Interestingly, criticism and accusations from the mayor of Sumac may have strengthened the district-union relationship, as both sides were fighting against the same criticism, which they perceived to be unfair. The mayor had long been critical of the school district and its teachers, which district stakeholders found upsetting. While public officials lambasted the district for having a low graduation rate, district leaders noted that it was actually improving and had surpassed the national average. When the new superintendent was hired, she had an implicit mandate to address this criticism by improving student outcomes. The desire to make improvements and, as one interviewee said, "prove the district's critics wrong," brought stakeholders together and made them willing to consider change.

Stakeholders in Laurel and Juniper, higher-performing districts, were not faced with the same outside pressure felt in Sumac. Parents and teachers alike were generally satisfied with the status

quo. There was less of a desire to “rock the boat.” Stakeholders in Juniper described an attitude of “if it ain’t broke, don’t fix it,” and noted that it was difficult to justify changes to the teacher evaluation system absent any obvious or widespread problems with teacher quality.

A Carefully Structured and Cooperative Approach Facilitated Progress

A district’s approach to considering change had a lasting effect on both labor-management relationships and their ability to ultimately adopt new policies. Stakeholder actions in the first stage of the reform process can set them up for success or contribute to future challenges. All three districts made the decision to form joint labor-management committees, which would be tasked with designing policy options for the district and union to consider. In no district did the superintendent present a pre-determined teacher evaluation policy or demand specific changes at the bargaining table. A stakeholder in Laurel described how this approach was important, given a historical lack of trust:

I think years and years of mistrust, it was one of those things where they didn't want to budge on certain things in the contract. And so, to be able to have this joint committee come together and take a look at one of those sort of untouchable areas of the contract for years and years, because they didn't have that positive rapport, was important.

This approach – first forming a committee and giving teachers a voice – lowered the stakes for teachers and made the decision to consider change less threatening. Teachers understood that they would be able to have input and that the union would need to approve the committee’s policy proposal in contract negotiations.

The process of agreeing to consider change and forming a joint labor-management committee took longer in some districts than others. Both Laurel and Sumac had preliminary committees before forming a group that would be officially tasked with engaging in the work of designing a new teacher evaluation system. The absence of a timeline and concrete “next steps” may have slowed the process. Because teacher evaluations are inherently a difficult topic to discuss, it can be challenging to build momentum and easy to table discussions. One stakeholder in Laurel recalled that their first ad hoc meeting did not result in a concrete plan for moving forward. He said, “And so I think we met maybe one or two other times that year and it... It just seemed like we needed to have more of a plan.” The lack of a carefully structured approach served as a barrier to progress.

Timing Makes a Difference

Timing is an important influential factor that cannot always be controlled. Timing can be fortuitous and help move efforts forward or prove to be a barrier, as unanticipated events unfold. As already discussed, in all three districts, reform efforts were aligned with transitions in leadership, which seemed to be a natural time to consider change. Some stakeholders also noted that the financial circumstances in their districts were improving, which may have enabled them

to take on new initiatives. One stakeholder in Sumac described how they had weathered the financial crisis of the Great Recession together, which strengthened their labor-management relationship.

There were many policy changes underway when conversations about teacher evaluation began. In some ways, this facilitated progress by building momentum. Districts were already transitioning to the Common Core standards and implementing new standardized tests, which could require teachers to change their instructional practices. If expectations of teachers changed, then the method through which they are assessed would need to be addressed, as well. Stakeholders in Sumac described the changes in standards, along with the hiring of new district personnel as the “perfect storm” for collaboration. However, the confluence of several policy changes can also prove to be overwhelming. Some teachers felt that too many changes were being considered at one time.

The timing of district reform efforts with events unfolding in other districts had an impact. Media coverage of current events both enabled and hindered progress at this stage. While media and current events may have motivated stakeholders to act, they also contributed to some stakeholders’ skepticism of that work. Efforts to implement performance-based compensation and get rid of tenure in other districts made teachers wary of their own districts’ intentions. One stakeholder observed:

I do think that they [the union] were suspicious about all the other publicity that goes on around this topic. Around whether this is to identify teachers that you’re going to get rid of or incentive pay.

Teachers in the case study districts specifically mentioned paying attention to several districts that were leading the way or at the center of controversies, including Washington, D.C. and Denver, Colorado. Two stakeholders in Juniper explicitly mentioned Michelle Rhee’s work to implement a new evaluation and merit pay system (IMPACT) and to eliminate tenure in D.C., which was in the media spotlight. Teachers in the case study districts were also aware of efforts to weaken job protections for teachers in other states.

While the national conversation may have been influential, many stakeholders were particularly motivated by more local events. It was not unusual for individuals to have worked in more than one school district in the area, and their perceptions were often informed by their experiences in other districts. Additionally, some union leaders had been involved with work the state teachers unions were doing around teacher evaluation. One especially influential event in California was the *Los Angeles Times*’ decision to publish controversial value-added scores for elementary school teachers in the Los Angeles Unified School District (LAUSD) in August, 2010 and again in May, 2011. Stakeholders in all three districts mentioned this controversy. This, along with the coverage of the broader teacher evaluation debate, made teachers nervous and may have fed union skepticism of their own leaders’ intentions. It also may have served as a catalyst for paying attention to and being active in local teacher evaluation reform efforts. As one individual said:

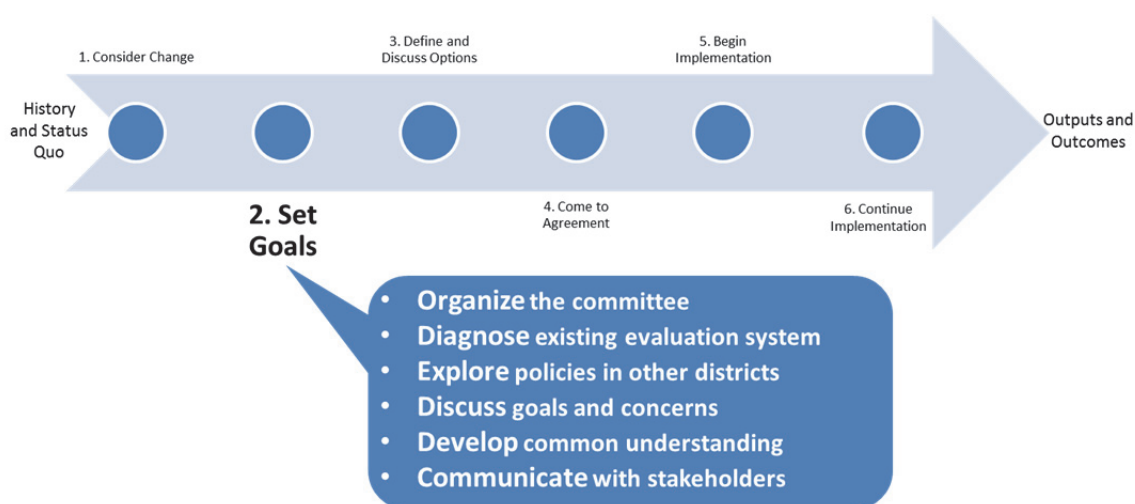
When we were first going out with this, was when the *LA Times* was printing all these scores, there was nervousness with that. I think, there again, it goes back to a general distrust between, particularly, union leaders and the administration.

Watching these events unfold in other districts made teachers in the case study districts feel that the stakes were high. For some individuals, this media attention heightened pre-existing distrust and may have made them less willing to consider change. While district and union leaders were able to get the reform process started, the underlying currents of skepticism remained.

Chapter Eight: Approach to Reform

In this section, I describe how the case study districts started the work of revising their teacher evaluation policies. Once the districts' evaluation committees convened, they engaged in goal-setting, the second stage in the reform process, which is summarized in Figure 8.1 below. This stage set the tone for each district's approach to reform. It is also where the important and difficult work of relationship-building began. In the discussion below, I first summarize the key activities and accomplishments during the goal-setting stage. Next, I describe the roles played by the various stakeholders and the character of their relationships. I then present the most salient factors that influenced their work.

Figure 8.1 Reform Activities – Approach to Reform



All case study districts formed joint labor-management committees to spearhead their teacher evaluation reform efforts. The goal-setting stage began when these committees first convened. (See Appendix G for a detailed comparison of the structure of each district's committee, across several characteristics, including leadership, membership, and accomplishments.) The composition of all three districts' committees' was similar and they consisted of eight to 22 individuals.²¹ The case study districts focused on several tasks during this second stage in the reform process (see Figure 8.1). There was variation in the amount of time each district spent on the various tasks, as well as the approach they took. As teacher evaluation committees began to

²¹ Sumac had the largest committee, though it is unclear if all 22 people who were involved at some point attended at the same time and with the same regularity.

meet, they were also engaged in the work of relationship-building, which was tacitly integrated into all of their activities.

Once regular committee meetings began, the groups organized by selecting a facilitator, creating a meeting schedule, and setting expectations. In all three districts, committee members spent time discussing their existing policies, identifying strengths and weaknesses. Committees then looked to research on teacher evaluation and explored evaluation tools from other school districts. Laurel's committee spent more time on research than the other two districts, while Sumac took a different approach and engaged in activities primarily geared at building trust. Committee members in all districts discussed goals for and concerns about teacher evaluation. The case study districts communicated their committees' work with other district stakeholders through various means.

After much discussion, the committees developed a common understanding of what they wanted to accomplish. Though they approached the goal-setting stage in different ways, the evaluation committees in all three districts came up with similar goals for their new teacher evaluation system. Table 8.1 summarizes the key activities during the goal-setting stage and the eventual goals established by each district's teacher evaluation committee.

Table 8.1 Committee Activities and Established Goals

| | Laurel | Juniper | Sumac |
|----------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Purpose of the committee | <ul style="list-style-type: none"> • Set goals, but not define new system • Submit letter of recommendations | <ul style="list-style-type: none"> • Develop a new system • Pilot new teacher evaluation | <ul style="list-style-type: none"> • Develop a new system • Pilot new teacher evaluation |
| Length of time in "Set Goals" stage | <ul style="list-style-type: none"> • One school-year | <ul style="list-style-type: none"> • Half of a school-year | <ul style="list-style-type: none"> • One school-year |
| Key activities of committees at "Set Goals" stage | <ul style="list-style-type: none"> • Diagnose existing system • Compare evaluation systems in other districts • Discuss goals • Post meeting notes and documents to district document sharing website | <ul style="list-style-type: none"> • Diagnose existing system • Visit a peer/model district • Review some research • Discuss purpose of evaluation | <ul style="list-style-type: none"> • Diagnose existing system • Define effective teaching • Develop guiding principles/goals • Compile resources to use in next stage • Create webpage to communicate with community, solicit input • Visit school sites |
| Goals established at this stage | <ul style="list-style-type: none"> • Create professional growth tool to support teachers • Update standards, use Danielson framework • Use multiple measures, but no VAM/SGM • Add beyond satisfactory level to ratings • Increase the frequency of observations • Incorporate peer review | <ul style="list-style-type: none"> • Create professional growth tool to support teachers • Move to standards-based system • Develop own standards • No VAM/SGM | <ul style="list-style-type: none"> • Create professional growth tool to support teachers • Move to standards-based system • Develop own standards • Include multiple sources of evidence • No VAM/SGM |

As shown in Table 8.1, all three districts came to the agreement that teacher evaluation should be a tool for professional growth, and that it should not be used as a “gotcha” for catching teachers doing something wrong. Committees emphasized the importance of using evaluation as an opportunity for coaching and as a means to improve teacher practice. Additionally, committees in all three districts agreed that value-added measures or other student achievement data would not be used to determine a teacher’s rating. Both Juniper and Sumac decided to move to a standards-based evaluation system, but did not want to adopt a set of standards or another district’s evaluation tool outright. They saw value in looking at other systems, but wanted a hand in designing the professional standards that would be used to evaluate their own teachers. Laurel already had an evaluation system based on the CSTPs, but committee members agreed that those standards needed to be updated. Stakeholders in Laurel were also very interested in having a rating category that would recognize teachers who go “beyond satisfactory.” However, in Laurel, the committee’s work focused more on setting goals for a new system rather than designing a new evaluation tool. In contrast, committees in Juniper and Sumac were tasked with the work of developing a teacher evaluation tool and pilot for the district.

Stakeholder Roles

There were many stakeholders involved in the goal-setting stage. As described previously, evaluation committees included district administrators, union representatives, and also school administrators. They took on the following roles:

- *Facilitators* who led committee discussions and organized the work;
- *Contributors* who actively participated in committees and completed assigned tasks;
- *Advisors* who occasionally attended committee meetings and provided support; and
- *Communicators* who updated and solicited input from district stakeholders.

The roles played by the key stakeholders in each district are summarized in Table 8.2.

Table 8.2 Stakeholder Roles in Goal-Setting

| | Laurel Unified | Juniper Unified | Sumac Unified |
|---------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Superintendent | Advisor Occasionally attended committee meetings. Listened to discussions and expressed support. Stayed updated on the committee's work. | Facilitator Led the committee. Tasks included setting the agenda, assigning work, facilitating discussion, and bringing in research and resources. Organized trip to a peer district. | Advisor Set expectations and checked-in with the committee from time to time. Communicated that this work was a priority. Received updates from the AS of HR. Communicator Gave updates to the school board and included relevant information in weekly memos. |
| Assistant Superintendent of Human Resources | Facilitator Shared the task of convening meetings facilitating discussions. Did research and presented to the group. Reported on progress to the superintendent. | Contributors Previous AS of HR served on the committee and participated in discussions. | Facilitator Previous AS of HR co-lead the committee until he left the district. Provided updates to the superintendent. |
| Other district admin | Contributors Served on the committee. Participated in discussions. One was the designated note-taker. Some took on research tasks and completed assignments between meetings. | | Contributors Served on the committee and participated in discussions. |
| Union President | | Facilitator UP#1 co-lead the first committee with the superintendent. Served as the facilitator of the teachers. Reviewed notes for accuracy. | Facilitator/ Communicator Current UP* served as the co-chair with the AS of HR. Did a site visit "blitz" to solicit member concerns. |
| Union reps / teachers | Facilitator/ Communicator One rep volunteered to lead the committee. Created agendas, assigned tasks, conducted research, and facilitated discussion. Communicated work to negotiations team/union reps. Contributors Served on the committee. Participated in discussions. Took on research tasks and completed assignments between meetings. | Contributors Served on the committee and participated in discussions. Some brought in resources for the committee to read and discuss. | Facilitator/ Communicator One union rep also served as a leader of the committee. Participated in a site visit "blitz" to solicit member concerns. Contributors Served on the committee and participated in discussions. |
| School admin | Contributors Served on the committee. Participated in discussions. Some took on research tasks and completed assignments between meetings. | Contributors Served on the committee and participated in discussions. | Contributors Served on the committee and participated in discussions. |
| School board members | n/a | n/a | n/a |

*Note: The current union president in Juniper was a representative when the committee started.

As shown in the table, school board members were not actively involved at this stage in any of the districts and that may have been intentional. One district stakeholder noted that it was important to keep school board members in their traditional role:

I think we work really hard to keep the school board in their proper role, which is around helping us with direction and vision, and then letting us do the work. And so, I think this is an example of where they definitely wanted to let us know what their interest was, and they did that, and then we set about to do the work and update them on how that work was going.

One school board member explained why he thought school board members should not be in the position of ultimately choosing an evaluation tool:

We're not educators and for the idea that somehow we would pick a model even though they're out there and have some really strong sense for what's better for educators. I don't think any of us believe that that's really our role or that we even have that capability.

Another school board member commented that they had other competing priorities that needed attention. The district superintendents provided school board members updates on the evaluation committees' work, but the degree to which each member paid attention was unclear.

Facilitators and Communicators

All district evaluation committees were jointly led by a representative from the district – either the superintendent or an HR administrator – and the union – either the president or a designated representative. At the goal-setting stage, the leaders took on the primary role of “facilitator.” Leaders convened meetings, set the agenda, assigned tasks, brought in resources and led discussions. They also took on tasks similar to those assigned to other committee members, such as researching the evaluation system in another district and presenting to the group, and consistently attended meetings.

In Laurel, the committee was facilitated by a union representative and the Assistant Superintendent of HR, with the union representative taking the lead. However, unlike in the other two districts, the president of the teachers union, who was on a leave-of-absence for part of the year, did not serve on the committee. In Juniper, the committee was co-led by the superintendent and the union president; however, interviewees considered the superintendent to be the primary facilitator of the committee. In the other two districts, the superintendents stepped back at this stage and delegated responsibilities to their Assistant Superintendents of HR. Interviewees in Sumac identified several individuals who filled the facilitator role at this stage. The previous Assistant Superintendent of HR and the current union president, who was a union representative at the time the committee began meeting, co-led the committee. An additional union representative also took on a leadership role. It was hard to determine if one of those individuals took on a greater leadership role than the others.

In Laurel and Sumac, several of the facilitators assumed the dual role of “communicator,” serving as spokespeople for the committees’ work and liaisons to stakeholders who did not serve on the committees (see Table 8.2). Because the superintendents in Laurel and Sumac did not attend all committee meetings, they were kept abreast of the work by their representatives, usually the Assistant Superintendent of HR. In Laurel and Sumac, union leaders were responsible for updating their executive boards and the entire union membership on the work of the committee. In Juniper, the communications task was less explicit. The two union facilitators in Sumac coordinated visits to all school sites to solicit concerns from their members. The superintendent in Sumac included updates about the committees’ work in her regular communications with the school board.

Contributors

Committee members who were not leaders contributed in a variety of ways. A committee member’s primary role was to participate in group discussions at meetings. As one Sumac stakeholder said, “The role of a committee member was to put your two cents in.” Members’ level of involvement at the goal-setting stage varied from consistent to occasional. In addition to engaging in discussion, some committee members contributed by taking notes and completing readings. In Laurel, individuals signed-up to take responsibility for researching other districts and then presented their findings to the group. Both teachers and administrators took on those roles.

School administrators, who were less involved in the considering change stage, became important contributors on the evaluation committees. Many stakeholders noted that it was helpful to hear the perspectives of individuals who had experience with conducting teacher evaluations under the districts’ pre-existing policies. Some district administrators had previously been school administrators and contributed this perspective as well. Several stakeholders commented that it was important to hear the perspectives of both teachers and administrators. Interviewees also found it helpful to have representatives with different areas of expertise and to have individuals who had participated in negotiations and were familiar with the contract language. As one Human Resources administrator said:

I was also the contract interpreter for the district so I would be able to kind of espouse or share some of the district's views on how pieces of the contract impacted our discussion, as did many other people at the table as well on both sides.

Advisors

In Laurel and Sumac, the superintendents served as advisors during the goal-setting stage. While they were invested in the teacher evaluation work and monitored it through their representatives on the committee, they did personally not take on leadership roles in the committees. In contrast, the superintendent in Juniper continued to spearhead the district’s

reform efforts at this stage. In Laurel, the superintendent occasionally attended committee meetings. When she did, she focused on listening to what committee members had to say, rather than advocating for a particular view. One committee member remarked, “I think she was there to listen more than she was to advise.” The superintendent thanked committee members for their participation and emphasized the importance of their work. Similarly, in Sumac, the superintendent’s role was to set expectations and then check-in on the committee’s progress periodically. She attended a meeting at the beginning, middle, and end of the reform process and likely attended only once during the goal-setting stage. She also made sure to communicate to the committee that their work was a district priority.

Stakeholder Relationships

The decision to form joint labor-management committees to address teacher evaluation policies set the stage for a potentially collaborative approach to reform in all three districts. Though labor-management interactions were not always positive, there were no serious conflicts that threatened to derail the process. During the goal-setting stage, committee members engaged in difficult discussions and expressed differences in opinion. However, having disagreements was not an indicator of poor relations. In fact, figuring out how to work through disagreements was an important part of relationship-building at this stage and in the following stages. Interviewees in all three districts noted that relationships were central to this work, not only in the context of decision-making, but in the eventual implementation of a new evaluation system. Stakeholders envisioned an effective evaluation tool as one that would foster coaching relationships between administrators and teachers, while focusing on professional growth. All committees agreed that building those relationships would be challenging if evaluation systems were primarily designed to be punitive.

During the goal-setting stage, the teacher evaluation committees had not yet begun the work of designing and debating concrete policy options. The lowered the stakes at this stage may have enabled the committees to amiably resolve disagreements. Again, Laurel seemed to have the most collaborative labor-management relationship at this stage, when compared to the other case study districts. Stakeholders described committee meetings a safe space for expressing opinions. Interviewees noted that committee members were willing to have dialogue. One individual commented on why they were able to have that dialogue:

It was never confrontational because it was... We were really engaged in an intellectual exercise. Not that we didn't believe it would go somewhere, but we were at that stage where you get to be fun thinking about it.

Committee members were able to discuss some contentious topics, such as the role of student achievement data. However, members were hesitant to challenge each other and may not have entirely resolved disagreements around sensitive issues.

A committee member in Laurel described the group as “collaborative, cooperative, productive, [and] positive.” Committee members were respectful and interactions were collegial. Members commented that there were no clear distinctions drawn between the role of a district representative and the role of a union representative in the context of the committee. As one individual described:

Well, we don't necessarily think about it as the association in the district. We were colleagues in the room. That's really how it was and how we felt.

Several factors may have contributed to this level of collaboration. First, Laurel made collaboration a district-wide goal for that year. Second, committee members in Laurel faced lower stakes for their work than other districts’ committees, as that group was not responsible for crafting an actual evaluation tool or making policy decision. Third, over the course of the year, they developed strong personal relationships. Interviewees described how members got to know each other personally, sometimes interacting outside of committee meetings, as well. One individual emphasized the importance of developing that relationship with the superintendent, saying, “You know each other. You walk in, you’ll see the superintendent, ‘Hey, what’s up?’ It’s kind of nice, right?” Interviewees considered that to be reflective of their unique district culture.

Fourth, district leadership in Laurel did not make demands of the committee; instead district leaders allowed the committee to guide itself, were open to criticism, and took on an active listening role. One committee member described feeling comfortable with offering her opinions:

I felt extremely comfortable being critical of current systems, current documents. Yes, for me to be able to be open and critical in front of the assistant superintendent, superintendent, I felt like they appreciated that honesty, but they don’t maybe hear that enough, even though it’s critical.

The administrators seemed to take more of a listening role, allowing the teachers to guide the debate. Fifth, the committee entered the process with consensus that the current evaluation system was broken. As a result, the committee members did not need to debate whether or not changes were needed.

Sixth, in Laurel, teachers did not express a concern that the district was looking for a punitive system. As one individual said:

And I think the evaluation, the vibe I got at every evaluation study meeting was we are not here to catch teachers in wrongdoing or in whatever, but it’s about growth. It’s about professional growth. It’s about instruction, what’s good for kids.

In contrast, concern about the true purpose of the evaluation system was a persistent issue in Juniper. Teachers, in particular, were concerned about the ultimate purpose of a new evaluation system. One stakeholder in Juniper said that concerns stemmed from the ambiguity in teacher evaluations and thought that “ambiguity can be used against teachers.” The district was unable to assuage some teachers’ concerns about the district’s intentions and took a defensive stance, which would prove particularly problematic later in the reform process. However, committee

discussions did not get contentious at this stage because the group did not yet need to vote on a specific policy option.

Juniper did move through the goal-setting stage more quickly than the other two districts. The labor-management interactions continued to be civil and somewhat cooperative. Though their interactions were shaped by the pre-existing climate of distrust, committee members did not act as adversaries. The superintendent, while taking the lead, recognized that she could not be the sole leader. Similar to the experience in Laurel, personal relationships helped the process. The superintendent and UP #1 had a friendly and collegial relationship and made an effort to stay on the same page throughout the process. They went back and forth with the notes from each meeting, making sure they both agreed that all points discussed were included.

As discussed, the committee in Sumac took a different approach to goal-setting than Juniper and Laurel did, making it an explicit task. They spent a year developing “guiding principles,” which were adopted by the committee by consensus. It was a challenging task and took significant time. The superintendent was new to the district and was still building relationships. People were described as guarded in the first meetings, though they gradually found ways to work well together. One stakeholder noted even observing a change in participants’ body language:

So I did know at the beginning, like I said, people came in with their own ideas and I saw, like I said, individual stance and even in body language, but they continued to have the conversation and talk and reflect and look at data and examining, until they got to the place where they were united.

Committee members in Sumac said that their comments were valued equally. As noted in Laurel, interviewees felt that there was no clear distinction between administrators and teachers. In fact, one individual commented on how they sat interspersed:

But when people are saying logical things, that doesn't matter if you're a this or a that. So, it really wasn't sides. Everyone sat interspersed, everyone was listened to equally, and like I said, everything that we did was through consensus. So, if one person was saying, "I've got an issue with this," you got to keep at it until you get to where everybody's got their thumbs up.

The labor-management interactions during committee meetings in Sumac were described as positive and productive. The committee released joint-communications to keep stakeholders updated, which helped build trust and keep everyone on the same page. One stakeholder said that people outside of the committee were surprised to hear how well it was going:

That began to dissipate the more these positive communications came out of the progress we were making, to the point where veteran people started saying, "This is great, the union and the district are collaborating on this, they're working together on this. The district didn't just go to the table and demand these changes, are actually facilitating discussion."

As the work continued, the committee itself may have had a positive effect on the labor-management relationship in the district as a whole. The extra time invested in the goal-setting stage laid the foundation for a positive relationship.

Influential Factors

While all three districts completed similar tasks at the goal-setting stage, they did not take the same approach, nor did they move at the same pace. The character of labor-management interactions and relationships varied across the districts, though none would be described as “contentious.” While Juniper and Sumac continued on from the goal-setting stage to begin drafting new teacher evaluation policies, the work in Laurel terminated with the committee’s letter of recommendations drafted at this stage. In the discussion below, I identify five factors that help explain what enabled districts to get their committees started and complete this second stage, and also discuss what made the work difficult.

The Role of the Superintendent

Superintendents continued to play an important role in the reform process, even when they were not directly leading the teacher evaluation committees. In Laurel and Sumac, the superintendents supported the reform efforts, but stepped back from the role of facilitator at the goal-setting stage. Their decisions to not serve as the committee facilitators seemed intentional. In fact, turning that responsibility over to other district and union leaders had a positive impact on the committees’ work in those districts; it allowed other stakeholders to take ownership of the initiative. Additionally, given that the superintendents of these districts were relatively new, it was important for them to work on establishing trust, which would be difficult to do while making immediate and sweeping policy changes.

Still, it was important for the superintendents to stay involved as advisors. By attending meetings occasionally, and particularly at the beginning of the process, they were able to clearly communicate their support to their committees. The superintendents’ endorsements were important. In Laurel, stakeholders knew that the superintendent wanted the committee to serve as a model of collaboration for the district as a whole. Having her at the table, while only occasionally, validated their efforts. One individual noted that she had participated in many district committees, and that they did not always have an impact:

Well, I think there are a gazillion task forces and committees, often formed because of a media situation or pressure from a particular constituency to look at an issue or solve an issue, and you're never really sure if your time is well-spent. Are you actually going to be able to make a difference? But when the superintendent is present, I think you have a greater sense of, "Oh, this is important work, and something will come of it."

Similarly, in Sumac, the superintendent communicated that she was committed to fostering labor-management collaboration. Some individuals noted that other committee participants adopted the superintendent's positive attitude over time.

In both Laurel and Sumac, committee members did not feel that their superintendent was advocating for a particular policy option. Participants saw them as open to their ideas and did not perceive the committee's work to be guided by a top-down mandate. That may not have been the case in Juniper, where the superintendent, who was the initiator of the reform efforts, continued to serve as the leader during the goal-setting stage. She was clearly dedicated to the work and wanted to ensure their efforts would be fruitful. The union president was also a co-facilitator of the committee, but the superintendent was seen as the primary leader. While it is impossible to determine if the superintendent's role impeded or propelled their work, it did not seem to improve labor-management relations. The existing climate of distrust in Juniper made relationship-building challenging. Likewise, it is not possible to conclude that outcomes in Laurel and Sumac would have been different if their superintendents took a more active leadership role; however, the decisions of their superintendents to step back had a positive effect on the committees' work in those districts.

Joint-Leadership and Broad Representation

All three districts decided to form joint labor-management committees outside of contract negotiations, rather than choosing a different approach, such as asking for specific changes to teacher evaluation policies at the bargaining table without previously consulting other stakeholders. Many stakeholders said it was common for their districts to form task forces around various issues, and taking that approach – making reform a joint effort – seemed particularly important in the context of teacher evaluations. One stakeholder commented that it may have been more challenging to tackle teacher evaluations in negotiations without first going through a committee:

I think if it came down in negotiation or something like that, it might get a little more contentious. But this was, to me, it seemed much more like a survey of what's out there.

The committee allowed teachers to play a leading role in designing a policy that would have a direct and potentially significant impact on their work and provided a forum to engage a diverse group of stakeholders. Many individuals explicitly noted that having broad representation was important and resulted in meaningful discussions. Teacher evaluation committees included individuals with different areas of expertise and experience. Interviewees noted that school administrators' involvement was particularly valuable.

A defining feature of a joint labor-management effort is having leadership from both sides. The willingness of district and union leaders to work together enabled them to convene committees. Leaders continued to be responsible for pushing the work forward and the relationship between district and union leadership continued to influence reform efforts in the

goal-setting stage. While all three districts had district and union representatives officially co-lead their committees, there were some differences in how those arrangements worked. In Laurel, the union facilitator seemed to serve as the primary lead, whereas in Juniper, the superintendent played that role. One individual in Laurel commented on why it was important for teachers to lead these efforts:

I think we let the teachers, the teachers really led it, because we really felt that especially if it's going to be something away from the table, we really wanted teachers to... Teachers had an interest in this, and so it really is important that if you're going to create something that's going to be meaningful for teachers, teachers need to be the ones leading that charge, I think.

In Laurel, committee members commented that they felt they were working toward a common goal. The teachers there did not express as much skepticism of the district's intentions with regard to teacher evaluations as they did in Juniper, where the superintendent took a more prominent leadership role. In Sumac, the leaders took turns facilitating. One committee member in Sumac commented on the benefits of that arrangement:

Neither one of them was elevated above the other, neither was more important than the other... And we weren't at the table like a bargaining table, where we're sitting on opposite sides. It was like the round table, where everyone had an equal voice in terms of what evaluation was about.

The committee in Sumac also released joint-communications to the district community, presenting a united front.

In addition to having effective joint-leaders, it was important for these leaders to ensure that the teacher evaluation committees included representatives from various stakeholder groups. In addition, the success of the committees relied on the inclusion of individuals who were motivated, open to change, and willing to collaborate. Individual participants, not just leaders, can have a significant influence on the committee's work, and leaders can influence participants' attitudes. Walton and McKersie (1991) found that authoritarian and hostile personalities can inhibit progress and collaboration. At this stage, committee members in all districts seemed committed to this work and motivated to move forward. Participants volunteered to take on various tasks and there were no individuals that actively tried to impede progress.

Topics On The Table

One approach to avoiding conflict was to keep certain controversial topics off the table. Generally, this was in regard to the use of student achievement data in teacher evaluation and related policies, such as merit pay. Districts made conscious decisions to not push for student achievement data or a particular set of professional standards. As one district administrator explained:

We thought that was a lightning rod point that would keep the system from moving forward, and so that was the decision we made.

It was clear from early on that including student achievement data as a measure of teacher effectiveness in a new evaluation system would be a non-starter for the union representatives on the committees. One committee member in Juniper said that the union representatives communicated that position at one of the first committee meetings. A stakeholder in Sumac recalled that some individuals expressed very strong opinions against using test scores. There was a particular distaste for value-added measures of teacher effectiveness, which may have been fueled, in part, by the *Los Angeles Times*' publication of teachers' scores.

As discussed, teachers in all districts were concerned about the purpose of a new evaluation system and districts maintained that the goal was not to punish teachers, but to support them. One school board member commented on their vision for the teacher evaluation reforms:

What we wanted to keep it out of was the whole space around, well, this is going to be used for terminations or for incentive comp or the whole value added discussion as well, even. Not that those things can't come into play down the road, but it really was, I think all of our vision that it was really focused on how do we help teachers work more effectively.

It was challenging for district administrators to convince committee members of their intentions. By removing the possibility that student achievement data would be a feature of their reforms, they lowered the stakes for teachers and narrowed the focus of future discussions. Limiting the scope of their discussions at the early stages served as an enabler of the committees' work as it helped facilitate open conversations. Student achievement data did come up in committee discussions, but in the context informing teacher practice or as one source of data that could drive coaching conversations. Maintaining a focus on professional growth was essential for keeping committee members invested and ensuring buy-in from stakeholders down the road.

Avoiding controversial topics may have helped districts avoid conflict among committee members, but that approach has potential downsides. Delaying challenging conversations is risky and it can complicate policy implementation or future quality improvement efforts. Though committees removed the largest contentious issues from the scope of discussion, committee members still had disagreements, and disagreements are not necessarily impediments. Navigating differences in opinion and learning to resolve conflict is an important part of relationship-building.

Stakeholders in Laurel reported a very positive environment, but also noted that committee members were reluctant to challenge each other. Though the group was collaborative, Laurel did not move beyond the goal-setting stage. Stakeholders in Laurel noted that teachers still had concerns about student achievement data that were never addressed. There was also no agreement on the possibility of a pay-for-performance system or eventually using evaluations to inform dismissal decisions, because those topics did not make it on the agenda. Some participants speculated that if they had actually been tasked with writing a new policy, the stakes would have been higher and individuals may have been more forthcoming with criticism. For Laurel to move forward with a new policy, these issues will eventually need to be resolved.

Time Invested In Goal-Setting

The goal-setting stage served several purposes. When committees first began meeting, not all participants knew each other. In addition, all three superintendents were new to their districts and did not have pre-existing relationships with committee members. The early work of the teacher evaluation committees provided the opportunities for participants to understand each other's points of view and get to know each other as individuals. It was important for the entire group to be working toward a common purpose. However, agreeing on common goals and building trust takes time. The committees in Laurel and Sumac both spent an entire school year engaging in this work. Juniper, which seemed to have more issues with trust, spent less time at this stage. The committee in Juniper did come to an agreement on general language of purpose, but may have benefitted from more time on setting goals. Time and patience are essential prerequisites for difficult conversations, and the time dedicated to discussing goals at this stage paid dividends later in the reform process.

The committee in Sumac took the most deliberate approach to defining common language. They developed a shared understanding of effective teaching, guiding principles, and clear objectives for the committee's work, and did not move forward until they achieved consensus. By engaging in that process as a group, it ensured that the reform efforts were motivated from the bottom-up. The group took this step slowly. One participant noted that it took a few months to get the committee established. By investing that time, individuals became comfortable sharing their opinions. One individual commented on how this set the committee up for success later on:

I think it's the process that we used. It wasn't top-down. Let's come together, figure out what our interests are, and see if this can serve both the district and the teachers, and we found that common ground. I think that's why it worked. I truly believe that's why it worked.

Taking time at the beginning, before jumping into debating policy options, allowed participants to build rapport with their colleagues. Participants noted that the time was well-spent. However, some individuals thought they could have moved more quickly. While relationship-building is a good investment, it is important for committees to make. As the case in Laurel shows, spending too much time at this stage may be inefficient and could slow the entire reform process down, particularly if there are not clear goals for moving forward.

Planning Ahead – Timelines and Clear Expectations

While dedicating time to goal-setting and relationship-building is important, having a structure that facilitates progress is essential for an effective committee. A lack of urgency, along with the absence of concrete next steps can derail the process. The district evaluation committees may have been more efficient at this stage if they invested time in planning ahead. One stakeholder commented on the importance of timelines:

I think timelines are critical. I think spelling out what consensus is, spelling out who the members are. All of that, I think is important.

A stakeholder in Sumac commented on how they would have benefitted from more structure at the beginning:

It was just like we were spinning our wheels and philosophizing. Oh my lord! The hours and hours of talking about what we believe and reading articles, establishing the common language. To me, it was overkill... I'm a little more task-oriented and I like to see structure and it would be nice if we were planning ahead and structuring out the meeting.

While the purpose of the committee in Laurel was to be a “study group,” one stakeholder reflected that he would have structured it better by starting with some “essential questions” that could provide direction. Their first few meetings were less organized and were mostly devoted to brainstorming. Laurel’s committee then decided to create a template for collecting and sharing their work, which helped direct their work. They established an agenda and decided on how they would run the meetings. Still, they did not realize how long this stage would take. As one administrator said:

I don't think people understood how long it would take to look at other systems. So as the year progressed there was a sense of where are we going with this? And we tried to. I think that toward the end of the year we were like, “Okay, let's pull this together.”

Some described the work in Laurel as “leisurely.” This could be due to the fact that they weren't motivated by a crisis or outside pressure. One individual noted how that could have contributed to the lack of urgency:

We didn't feel the sense of urgency that we've got to fix something that's so broken, we've got to get on top of this. Or there is all this heat out there, we've got to get control of it and if we don't, it's going to blow up at us.

Timing was also a barrier in Laurel because the work was not aligned with the negotiations cycle, which could have been addressed with better planning. By the time the committee released recommendations, negotiations for that year had concluded. It was too late to make any policy changes. While Sumac and Juniper moved on to the next stage, the work in Laurel stopped progressing. A focus on being collaborative and avoiding conflict, while conferring benefits, may have had some negative side effects. As one participant in Laurel said:

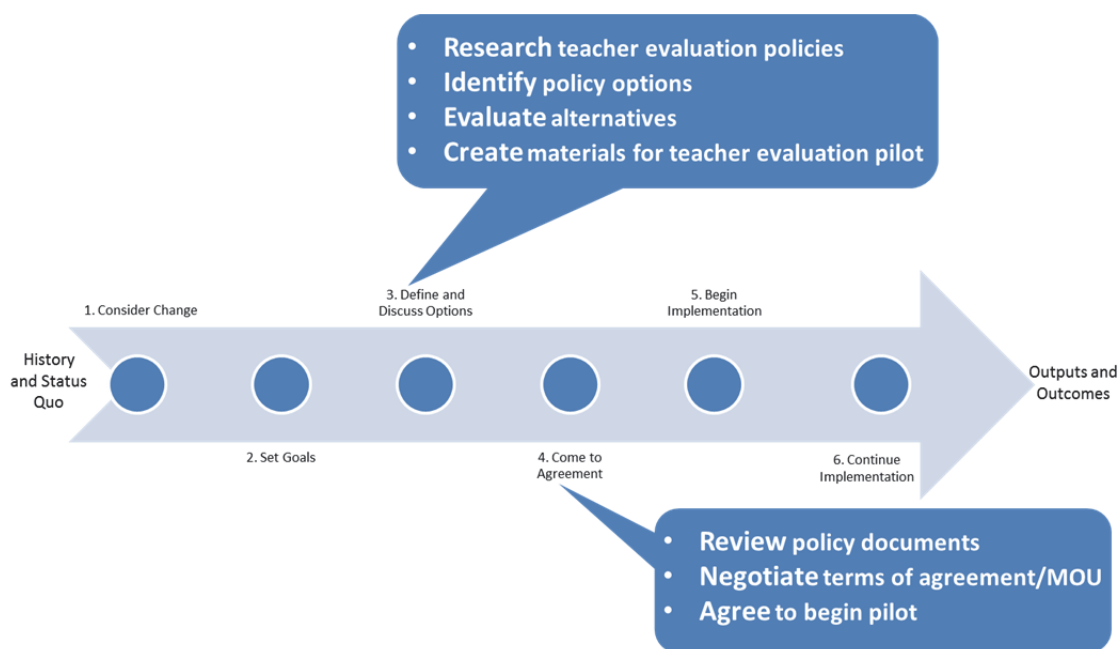
I think that one of the things that we sometimes, we are big on process and getting lots of input and voices at the table. We've been around a long time. You have to do it. If you don't do it, it's not going to work. I think the down side of that is, what happens many times is, we do very substantive, comprehensive, highly collaborative processes that yield very little results. Because by the time you're done you're so exhausted because it took so much out of you.

Finding a balance between taking time to build relationships and be inclusive, while also accomplishing objectives is difficult and requires careful planning.

Chapter Nine: Hammering Out the Details

While Laurel's evaluation committee stopped meeting after the goal-setting stage, the committees in Juniper and Sumac went on to hammer out the details of their new teacher evaluation policies. In this section, I review the work that Juniper and Sumac did in third and fourth stages of the reform process: define and discuss options for change; and come to agreement. The key reform activities observed in these stages are displayed in Figure 9.1 and summarized below. I examine how stakeholders' roles and relationships evolved at these stages and continue to explore the enablers of and barriers to progress, paying particular attention to the setbacks that Juniper experienced when the membership of its teacher evaluation committee changed.

Figure 9.1 Reform Activities – Hammering Out the Details



During the third reform stage, the teacher evaluation committees defined and discussed concrete policy options. They continued to look at examples of teacher evaluations from other districts and from their research; however, both Juniper and Sumac decided to develop their own evaluation tools, rather than adopting a pre-existing system, and then try them out as pilot programs for two years. Committee discussions spanned many topics and there were some major disagreements that needed to be resolved. The most challenging tasks were writing their own professional standards, deciding on the number of rating categories, and then designing a rubric that would distinguish between levels of effectiveness for each standard. Committees were

responsible for creating all the necessary documents for the new evaluation system, including policy descriptions and classroom observation forms.

In the previous reform stage, committees in both Juniper and Sumac agreed that they wanted to move to a standards-based teacher evaluation system. They then had to decide which set of professional standards to use as a model (the CSTPs or NBSTP), how many standards to include, and how many standards a teacher would be evaluated on at once. Stakeholders had greater reservations about standards that seemed to be the least objective. The most debate was around standards relating to professional practice, which were considered to be the most difficult to understand and to observe.

None of the case study districts intended to base teacher ratings on student achievement data, as was evident from committee conversations during the goal-setting stage. However, the role of student data did come up again as committees were hammering out the details of their evaluation policies. Student data is referenced in both the CSTPs and NBSTP and used by other districts that served as models. Both Juniper and Sumac came to the compromise that teachers should be using this data to improve their instruction and that administrators should also utilize this data when coaching teachers, but that student achievement measures should not be factored into their overall rating.

In their existing systems, all three case study districts had three final rating categories: satisfactory, needs improvement, and unsatisfactory. While there was agreement that the ratings needed to be expanded from the existing three levels, including a category that went beyond satisfactory, there was debate over what form that would take. Both districts also considered the role of peer review in teacher evaluation and the importance of allowing teachers to choose their own professional goals. After substantial discussion, both Juniper and Sumac reached stage 4, “come to agreement.” They successfully designed and agreed on new teacher evaluation systems to be piloted in the 2013-2014 and 2014-2015 school years. While Juniper’s pilot had to be approved in contract negotiations, Sumac was able to move forward without an MOU. Juniper did experience setbacks at this stage when the composition of its teacher evaluation committee changed.

Stakeholder Roles

Stakeholders’ roles on the teacher evaluation committees were similar to those in the goal-setting stage (see Table 8.2). There were facilitators, who led the committees’ work; members of the committees who contributed to the work; communicators, who engaged stakeholders outside of the committees; and advisors, who were committed to the work, while not being directly involved. However, there were some changes to committee membership, most significantly in Juniper. Also, the teacher contract negotiations team in Juniper, which included representatives not otherwise involved in the teacher evaluation reform efforts, played a role in crafting an agreement.

Facilitators and Communicators

Sumac's current union president, who was already serving as a leader of the committee in her position as a union representative, officially took office at the beginning of the 2012-2013 school year, when the committee began to define and discuss policy options. Her role as a facilitator on the committee remained the same. The current Assistant Superintendent of HR in Sumac joined the district in the middle of the school year, as the committee was discussing policy options. He stepped right into a leadership role, taking over from the previous individual in his position, and helped to facilitate committee meetings. That spring, committee documents officially referred to the union president and new Assistant Superintendent of HR as co-chairs. However, interviewees considered two other individuals – a union representative and another district HR administrator – to be facilitators of the work, as well, because they took on additional responsibilities and helped drive the process forward. The union president and Assistant Superintendent of HR also assumed the responsibilities of communicators, presenting at school board meetings and ensuring that district stakeholders received updates after every committee meeting. The committee created an online survey that gave all district stakeholders the opportunity to offer their opinions on the teacher evaluation work before the terms of the pilot were settled.

In Juniper, the superintendent maintained her role as facilitator as the committee debated options and designed the terms of their agreement. She was also involved in contract negotiations, where the MOU that approved the pilot was drafted. The second union president took office during this stage and was also involved in negotiations, though she would not serve as a leader on the teacher evaluation committee. The MOU explicitly stated that the second teacher evaluation committee would be jointly chaired by the union and district. The superintendent and a union representative, who had served on the first committee, assumed the roles of co-chairs.

Participants

Over the course of the year, there were around 20 participants in Sumac's teacher evaluation committee meetings. When they released their final product to the district community, 22 individuals were named as committee members, including district administrators, school administrators, and teachers union representatives. Given how many individuals were involved and the number of tasks that needed to be accomplished, Sumac formed several sub-committees at this stage. Participants took on additional responsibilities for tasks as members of those sub-committees. For example, one sub-committee was responsible for creating a working draft of all the evaluation documents to present to the larger group.

Juniper's committees were smaller, with only eight members. UP #2 appointed two new union representatives to sit on the second committee; the previous union president was not invited to return. The superintendent also added new district representatives. While the role of a committee participant in Juniper did not change, the new committee participants did raise their

own concerns and objections. Though the new members did not try to derail the process altogether, they did cause the work to backtrack and the superintendent faced more resistance from this group.

While there were only eight members on each of Juniper's committees, the individuals representing the district and the union during contract negotiations were also important participants at this stage as they were tasked with reviewing the documents created by that committee and crafting the MOU that would authorize the pilot.

Advisors and Observers

Sumac's superintendent remained in an advisory role and was not an official member of the committee. She did communicate to the new Assistant Superintendent of HR that they needed to make progress on the teacher evaluation work and come to an agreement on the pilot; she expected him to push that work forward. As the committee continued to engage the district community, I assume that more individuals became "observers" of the teacher evaluation work, though I cannot identify them all by name. In Juniper, the teacher evaluation work garnered more during the middle reform stages, from school board members and others, because it was a topic addressed in the very contentious contract negotiations. The school board followed the negotiations proceedings closely.

Stakeholder Relationships

Labor-management interactions looked different in Juniper and Sumac during the middle two stages of the reform process. Both teacher evaluation committees engaged in some tough conversations, as is to be expected when defining and debating policy options; however, while relationships in Sumac continued to improve, the labor-management climate in Juniper deteriorated.

In Sumac, both the district and union representatives who were interviewed agreed that interactions among committee members were positive during these stages. Some of their disagreements had already been resolved in the creation of their guiding principles, which they used as a point of reference during debates. Interviewees described committee members as passionate about the work and viewed committee meetings as a safe space to engage in difficult conversations. The group stayed committed to working through tough issues and finding consensus:

We used consensus as our operating mode so I know there were like really hard conversations we had but when you truly are trying to use consensus in getting to yes, you don't let those things get in your way.

This mirrors the IBB approach that Sumac uses during contract negotiations.

In the third and fourth reform stages, the labor-management relationship in Sumac fit the description of "collaborative." Participants remained diplomatic and polite throughout the

process. By the end of the year, when an agreement had been reached, committee members had built trust and worked together to meet common goals. As one individual described:

In the very beginning, it was more like hard rocks, individual hard rocks, you know? And then at the end I saw this beautiful marble that had kind of melted together and around student achievement.

Another person noted how they were able to build trust, saying, "I just saw little by little any misgivings or mistrust begin to melt away. And I still think that's paying dividends." Having the new Assistant Superintendent of HR assume the role of co-chair right after he joined the district could have stirred controversy; however, there was no resistance. The committee members had already built strong relationships, and they reacted positively to the new leadership. As one individual said:

And so in January, he picked up as helping to co-chair that committee and it continued on without a hitch. It was like, "Okay, we got the new coach here." But because the committee itself was so strong and was speaking with one voice and doing things so collaboratively, it just... It could run itself. But obviously, he... It was awesome with him coming in, because he brought in his perspective and seeing things from his side, which was great. So he was one more voice, in terms of everything we were doing.

Stakeholders saw the continued turnover in district leadership, led by the superintendent, as having a positive impact on the district. They trusted the superintendent and were pleased with her hiring decisions.

The relationship-building that happened in the context of Sumac's committee helped cultivate trust between the district and union outside of the committee, as well. The committee regularly communicated their work to the entire district community, rather than acting in isolation, which one individual noted was essential to building trust:

So we made sure that the school board knew what was going on. We made sure all of our 1600 plus [teachers union] members knew what was going on. That all the administrators knew what was going on and I think that's one of the best things, transparency. Making sure that you're reporting regularly to everyone, where everyone gets the same information at the same time. And people are able to give their input and then we say, "Hey, we've considered this input and we've done this or that, or we stayed with what we have because we believe blah, blah, blah." All that builds trust.

As previously noted, the committee was able to come to an agreement outside of contract negotiations at this stage. The fact that they were able to begin the pilot without an MOU was a testament to the relationships they cultivated. As one district stakeholder said:

And if we did not have a good rapport with the unions, I still think about what we did without an MOU and I go, "Wow, we really stretched it."

They maintained a collaborative relationship as they moved on to the implementation stage.

In Juniper, negative perceptions of the district's intentions continued to hamper the committee's work during the discussion and agreement stages. In the first committee, though

trust was limited, interactions were mostly civil. When the second committee was formed under the new union leadership, labor-management interactions changed. The superintendent did not have the same relationship with the second union president as she did with the first. UP #1 had worked closely with the superintendent and endorsed several district initiatives. UP #2 had not previously worked with the superintendent and was not as willing to co-sponsor initiatives with the district after taking office. Some district and union representatives speculated that the new union leaders changed the composition of the committee because they saw the group as being too collaborative and did not want to retain members appointed by previous union president. Even the way that committee members communicated was affected by the change in leadership. As one stakeholder commented:

Communication has become very difficult. Whereas communications could be informal and productive before, now they are always rather formal and combative.

Interactions between the superintendent and the new union representative were at best civil and sometimes contentious, especially outside of the committee.

Interviewees familiar with both committees noted that the second committee expressed less enthusiasm about the teacher evaluation reform efforts and though they agreed to move forward with the pilot, they seemed less invested in the work. Some union representatives viewed the pilot as the district's own initiative, rather than a joint-venture. One individual described the agreement reached in negotiations as a "barter," rather than a plan reached by consensus:

Because political interest drove it. Should I tell you that we got a raise in return? We gave up language that is really shoddy language that doesn't do the job... It was a barter. I'm not quite sure that an evaluation system should be created on a barter. You would think that there's... "I really think this is shitty, but if you pay me, then I'll live with it."

It seemed that the union only accepted the new teacher evaluation pilot because they received the other concessions from the district, and that they did not embrace the new system. Interviewees noted that the agreement was not viewed as collaborative, as one individual described:

It's not collaborative. I mean collaborative would assume that both sides consider the other side's point of view. And from our perspective, and I think it's true for both of us, both sides, that from our perspective, it was more like strong arming the other side to recognize... I mean it's just, it's not collaborative. You could label it anything you want, but at the end of the day when the only way I am getting stuff done is for you to not be able to move forward and for you to get things done, is for me to not be able to stand my ground. It's not a collaborative effort.

Outside of the committee, the labor-management interactions were particularly contentious. Contract negotiations were described as "tense," and the lack of trust between the district and the union was apparent. Significant conflict over the certificated salary schedule emerged and the

union organized public protests. This overall antagonistic climate affected the work of the committee.

Influential Factors

Many of the factors that influenced reform efforts at the beginning stages of the reform process continued to be relevant as the districts hammered out the details of their teacher evaluation pilots. Superintendents, and other stakeholders' perceptions of them, continued to have a significant influence on labor-management relationships. The roles of the superintendents in Juniper and Sumac did not change much as the evaluation committees moved to the decision-making stages. Leadership, in general, remained a key factor in shaping the reform process, and both districts experienced turnover in leadership at these stages.

Committee structure proved to be important in the middle stages, both in terms of membership and approach to the committees' work. The time invested in the goal-setting stage began to pay dividends for Sumac and may explain why Sumac had more positive relationships at this stage. The topics of discussion and reform features that stakeholders advocated for influenced the committees' ability to reach agreement and shaped the compromises that they made. District culture colored the committee's work and at the same time, the committees were shaping labor-management relations in their districts. In Juniper, a conflict between the district and teachers union during contract negotiations was an obstacle to moving the teacher evaluation work forward. In this section I discuss the key factors that enabled districts to successfully design a pilot and what challenges they had to overcome to reach agreement.

Leadership Turnover Can Pose Challenges

While changes in leadership may have propelled districts to consider change in the first place, leadership turnover during the reform process posed challenges, particularly in Juniper. In that district, the change in union leadership made collaboration challenging. The first union president left the teacher evaluation committee and the new union president and superintendent did not have a pre-existing working relationship. Instead, their relationship was being shaped by the concurrent and contentious contract negotiations. When the second evaluation committee convened with new members, communication became more difficult and previously-settled disagreements were resurrected. While the ensuing debates did not delay the implementation of the teacher evaluation pilot, they did put a strain on labor-management relations. This also affected the manner in which committee decisions were reached. As previously discussed, teachers on the second committee equated "coming to an agreement" with "giving in" to the district. One stakeholder described it as "miraculous" that, after all that conflict, they were able to move forward with a pilot.

Committee leadership turnover during the middle reform stages had different effects in Sumac than it did in Juniper. The addition of the new Assistant Superintendent of HR to the

teacher evaluation committee did not harm the labor-management relationships; the trust that was previously established over the course of the committee's work was not compromised. The new district representative was viewed as an extension of the superintendent and union representatives trusted his judgment. Though a new union president was elected, she was already serving as a committee co-chair. As a result, leadership shifts did not result in the same loss of continuity that Juniper faced. Relationships developed in Sumac's long goal-setting stage may have buffered the leadership transition. The guiding principles they established remained in place even when new committee members joined and the committee members' ability to work together improved over time.

District Culture and Contract Negotiations

While Sumac was able to come to agreement on an evaluation pilot outside of contract negotiations, Juniper sought an official MOU before moving forward.²² One stakeholder in Sumac stated, "We wouldn't have been able to do that if the union and the district weren't comfortable and didn't trust each other." In Sumac, the teacher evaluation committee had a positive influence on the labor-management culture in the entire district. Their efforts to communicate their work and be transparent helped to build support for their work. The committee released joint-communications to the district community, which is not something that happened in Juniper. They also gave stakeholders the opportunity to provide feedback. People realized the district was not demanding changes, but they were actually facilitating a discussion. The committee itself served as a model of collaboration:

But I do think that that whole process did have an impact on our culture out there. People are like, "Wow, they can work together. It doesn't have to be combative all the time."

The establishment of trust would be crucial for successful implementation.

In contrast, in Juniper, the existing culture of district and the absence of a pre-existing relationship between the superintendent and newly-elected union president, made negotiations difficult. While the evaluation committee was separate from the district's and union's bargaining teams, some individuals were members of both, including the superintendent. The conflict during negotiations spilled-over into the teacher evaluation committee and exacerbated an already tense situation. The entire district community was aware of the arguments and accusations coming from the bargaining table. When the second teacher evaluation committee convened after negotiations were complete, the labor-management relationship was strained.

²² Teacher evaluation policies would later be addressed in contract negotiations at the end of the first year of the pilot.

Compromising on Reform Features

In an effort to avoid conflict, both districts continued to keep some controversial topics off the table, as they did in the goal-setting stage. The committees did not debate the use of student achievement data to calculate a teacher's rating; rather, they discussed alternative roles that student data may play in the overall process. Even though some district administrators wanted to see their districts utilize student achievement data further, they were willing to compromise and keep the focus on professional growth. Neither district explored the possibility of tying teacher evaluation to compensation or tenure decisions. However, not all stakeholders were convinced that these policies were completely off the table. One committee member in Juniper commented that some teachers remained concerned about what the new evaluation policy would evolve into:

There's a faction of people who were distrustful that the process wasn't... They thought that the process was just looking for a way to fire teachers, to get rid of teachers, although it had no effect on ed code or tenure; however, all you need to do is say that really loudly multiple times and people believe you. And so that was a concern.

The culture of distrust in Juniper may have made it difficult to assuage these concerns. Also, not addressing these topics fully during the goal-setting stage may have contributed to the lingering concerns about how the new system might be amended in the future, which some teachers still held. Even though student achievement data would not be included in the current revisions to the evaluation policy, the district may advocate for its inclusion in the future.

As described earlier, there was significant conflict around the topic of professional standards, especially those related to professional practice, which addressed areas some teachers felt were out of their control. The specific language used in the evaluation rubrics proved to be a stumbling block, especially in Juniper, which ended up with a much larger set of standards and sub-standards than Sumac.²³ Juniper, which adopted the NBSTP as a model for their evaluation, had two professional practice standards, whereas Sumac, which used the CSTPs, only had one. Juniper's second committee spent much time debating the language of those standards. They had to balance having enough detail, so that standards were not ambiguous, with having a manageable number of standards. Interviewees noted that disagreements over the language of the rubric were difficult to resolve in Juniper and that the prolonged debate was frustrating. Having fewer sub-standards may have helped Sumac curtail that discussion.

Both districts agreed to do a pilot before adopting a permanent system, which made the agreement less threatening, especially in Juniper. The union knew they could later re-negotiate the terms, and were thus willing to accept the changes to the evaluation policy. The pilots were also of limited scope. In Juniper, only probationary and temporary teachers were required to

²³ Juniper's six standards were divided into 63 sub-standards; Sumac defined only 32 sub-standards. However, since teachers in Juniper were only rated on four standards in any given evaluation cycle, they would be assessed on 34 to 38 sub-standards in one year, which was similar to the number of sub-standards in Sumac.

participate and in Sumac, only five schools were selected to participate. Both districts planned to roll-out their pilots over two years so that improvements could be made, if necessary.

Committee Structure, Planning, and Persistence

A persistent theme throughout all reform stages has been the importance of structure and timelines for moving forward. Without a concrete plan, research tasks and debates can continue indefinitely. Sumac had a very structured approach as they moved into the decision-making stages, which enabled their progress. The committee created an outline of discussion topics for their meetings for the entire year. They also decided to create sub-committees, which were responsible for specific tasks, and that made the reform process more efficient. It is unclear if Juniper would have benefitted from the same approach, as their committee was smaller than Sumac's. Having sub-groups of people take over certain tasks, independent of the larger committee, could actually serve as a source of conflict if there is not a strong foundation of trust in the larger group.

While both districts wanted to make progress, they also emphasized the importance of not rushing, as was also observed in the goal-setting stage. One stakeholder in Juniper said it was important to tackle one topic at a time:

You know, we just took small chunks at a time. We just work with it, and look at what here can't you live with, and the fact that we were willing to take certain things off the table, I think was helpful. Some things they absolutely agreed with. They were just... They wanted to take time, not rushing was very important, and just eventually coming to the end. By the end, I think we all wanted it to be over.

Stakeholders in Sumac also saw the benefits of not rushing through their work. They made a deliberate effort not to rush and to really think through options and how they would be implemented in their district. They commented that going slow in the early stages enabled them to find common ground and then come to agreement more quickly in later stages. One individual commented on the process of coming to agreement:

And I was shocked it was so fast. And, but I think, the old thing, "You've got to go slow to go fast". I think that really is apropos here, because that painstaking slow beginning seems to pay off later.

Stakeholders have to remain committed to their work, especially when it is approached slowly, as the work can become tedious. As one Juniper stakeholder said:

And so I think ultimately that, plus we stuck with it too. It would be a pretty easy thing to give up on if you didn't really continue to try to focus on it. With all the downturn in education and all the struggles for the last six years, it would have been easy to just say, "Let's do that down the road."

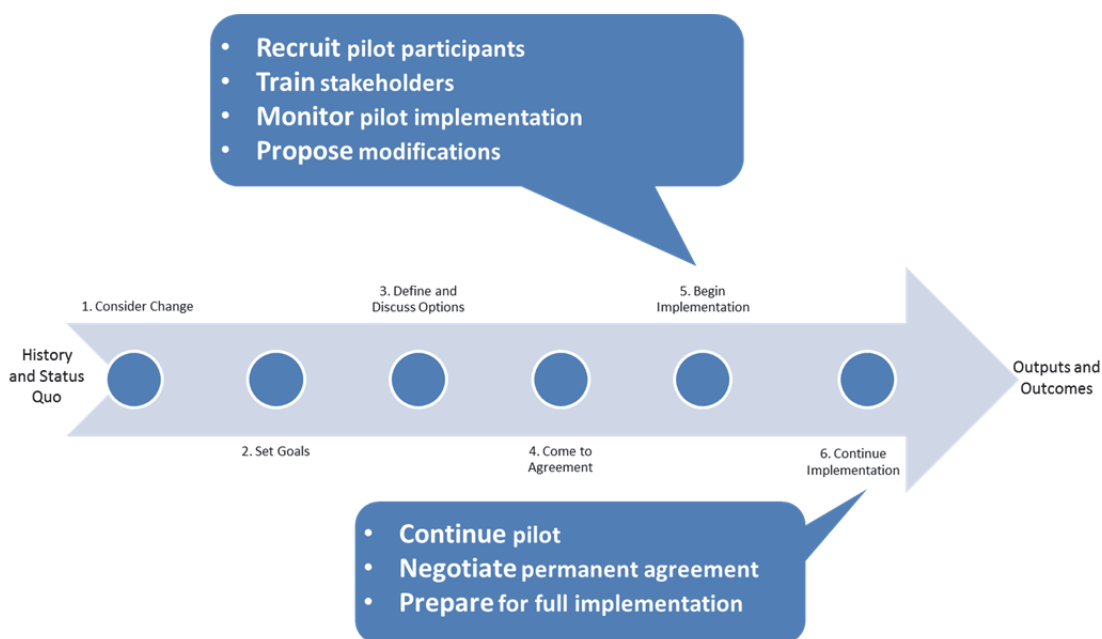
Both Juniper and Sumac experienced conflict and faced challenges, but persisted, led by individuals who remained committed to seeing a return on their investment. In contrast, Laurel

never reached these middle stages and, as previously discussed, the lack of structure to their efforts may have allowed the process to stall.

Chapter Ten: Moving Forward

This chapter summarizes the implementation stages observed in Juniper and Sumac and addresses the current state of the districts' teacher evaluation policies. I review the activities undertaken during first year of the teacher evaluation pilots (see Figure 10.1) and then comment on the transition to the second year of implementing the pilots.²⁴ As in previous chapters, I characterize stakeholder roles and relationships and examine factors that influenced early implementation.

Figure 10.1 Reform Milestones – Moving Forward



Implementation of the newly adopted teacher evaluation policies began in the fifth stage of the reform process. Both Juniper and Sumac embarked on two-year pilot programs, commencing in the 2013-2014 school year. In that first year, districts engaged in several activities. They provided training and ongoing support to pilot participants and other stakeholders. They solicited feedback from participants and held committee meetings to monitor progress. Committees proposed improvements to be considered before the second year of implementation, the sixth reform stage. Both districts made their revised teacher evaluation policies permanent during contract negotiations in 2014, with full implementation slated for the 2015-2016 school year.

²⁴ I only collected data on the first year of the pilot, but received updates from one representative from each district during the second year.

Table 10.1 outlines the characteristics of the teacher evaluation pilots in each district.

Table 10.1 Teacher Evaluation Pilots

| | Juniper | Sumac |
|---------------------------------------------|-------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Pilot Years | Year 1: 2013-2014 Year 2: 2014-2015 | Year 1: 2013-2014 Year 2: 2014-2015 |
| Full Implementation | 2015-2016 | 2015-2016 |
| Participating Teachers | Temporary and Probationary teachers required; permanent teachers must opt-in | Year 1: All teachers at 5 pilot schools Year 2: All teachers at additional pilot schools <i>Teachers could opt-out (1 high school teacher did)</i> |
| Number of Participants | Year 1: 25% of teachers (32% of participants were permanent teachers) Year 2: 31% of teachers | Year 1: 12% of schools Year 2: 75% of schools |
| Standards Adopted | 6 standards, based on NBTPs (Includes 63 sub-standards) | 6 standards, based on CSTPs (Includes 32 sub-standards) |
| Number of Standards Evaluated | 4 standards in one evaluation cycle | 6 standards for permanent teachers; 3 for temporary and probationary teachers |
| Goal-setting | Pre-evaluation form due in October, teacher and principal each pick one standard as a growth goal | Optional self-reflection form; teacher selects two goals on goal-setting form |
| Conferences | Goal-setting conference in the fall Lesson de-brief (option to bring an artifact) | Fall evaluation conference/goal-setting form Pre-observation conference (at least 1) Post-observation conferences (at least 3) Final evaluation conference |
| Observations | Temporary and probationary teachers: 2 Permanent teachers: 1 At least 30 minutes each | At least 1 scheduled (<60 minutes) At least 2 unscheduled (<30 minutes) |
| Observation Rubric Rating Categories | <i>Scale of 1-6</i> Exceeds standard (5-6) Meets standard (3-4) Does not meet standard (1-2) | <i>Scale of 1-4</i> Exceeds standard (4) Meets standard (3) Partially meets standard (2) Does not meet standard (1) |
| Final/Overall Rating Categories | Meets or exceeds standard Does not meet standard | Exceeds standard (4) Meets standard (3) Partially meets standard (2) Does not meet standard (1) |

Note: The reform features described in this table reflect what the policies as they were implemented in the first year of the teacher evaluation pilots.

Sumac selected five schools to participate in the first year of the pilot. Juniper, a smaller district took a different approach to recruiting for the pilot and did not end up with participants from every school. Only teachers who were due for an evaluation were eligible to participate in the pilots; the new policies did not change the frequency with which teachers were evaluated. While both districts adopted similar policies and implementation timelines, there were some notable differences. As shown in Table 10.1, Juniper adopted a set of teaching standards based on the NBSTPs, while Sumac used the CSTPs. Teachers in Sumac received a final rating in one of four effectiveness levels, whereas teachers in Juniper were rated as in “meets or exceeds standard” or

“does not meet standard.” Sumac required more teacher-administrator conferences and more classroom observations than Juniper did. Additionally, each observation, scheduled or unscheduled, required a follow-up conference. Both districts required a goal-setting conference with the teacher and administrator in the fall, and a post-observation conference to debrief the observed lesson.

In the first year of implementation, both districts made efforts to communicate the details of the teacher evaluation pilots with stakeholders and provided multiple training opportunities. Sumac conducted joint trainings for administrators and teachers participating in the pilot, while teachers and administrators were trained separately in Juniper. Since implementation began, Juniper and Sumac have engaged in ongoing efforts to monitor and improve their evaluation systems. In the first year of implementation, both districts solicited feedback from pilot participants in a variety of forums. The districts received both positive and negative feedback and both made changes to improve their systems. Both Juniper and Sumac intended to continue the quality improvement efforts they began in the first year of implementation during the second year of the pilot. The evaluation committees would continue to meet and consider improvements to their teacher evaluation policies, in anticipation of full implementation the following year.

Stakeholder Roles

At this first implementation stage, the number of stakeholders involved in the teacher evaluation reform process necessarily grew. In both Juniper and Sumac, pilot participants included principals and teachers who had been previously uninvolved in their district’s reform work. Individuals who participated by implementing the pilot at their school site will be referred to as “contributors.” New roles emerged at this stage to reflect stakeholders’ work outside of the committee setting. (In Juniper, the second teacher evaluation committee stopped meeting during the first year of implementation; while Sumac’s committee continued to meet.) I labelled the leadership role during the implementation stages as the “implementation lead,” referring to the individual primarily responsible for training pilot participants and monitoring the roll-out of the program. In both districts, this role was played by the Assistant Superintendent of HR, with significant support from a few other stakeholders, who will be referred to as “implementation support.” Other stakeholders remained involved as communicators, contributors, and advisors. School board members stayed tangentially involved as observers, who supported the districts’ work. In Juniper, school board members specifically requested to be trained in the system and, in the second year of the pilot, Juniper would add a board member “observer” to its third teacher evaluation committee. See Table 10.2 for an overview of the stakeholder roles in the first year of the teacher evaluation pilots.

Table 10.2 Stakeholder Roles in Implementation Year 1

| | Juniper Unified | Sumac Unified |
|---------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Superintendent | Implementation Support Presented at some school sites and to the school board. Conducted a demo lesson. Met with pilot participants to solicit feedback. | Advisor Less present at this stage. Checked-in with the committee from time to time and received updates. Supported the pilot. |
| Assistant Superintendent of Human Resources | Implementation Lead/Communicator Served as the point-person for the pilot. Presented pilot at all school sites. Led trainings. Conducted demo lessons and observations. Supported principals. Presented to the school board. Administered surveys. Met with pilot participants to solicit feedback. | Implementation Lead Served as the point-person for the pilot. Led trainings and info sessions. Conducted observations. Supported principals. Presented to the school board. Administered surveys. Met with pilot participants. Facilitator/Communicator Convened meetings and facilitated discussions. Communicated work with stakeholders. Reviewed feedback. Proposed and voted on policy changes. |
| Other district admin | <i>Served as committee members in the second year of the pilot.</i> | Contributors Served on the committee and participated in discussions. Reviewed feedback. Proposed and voted on policy changes. Implementation Support One HR administrator was very involved in supporting implementation and creating training materials, including a video for the evaluation website. Participated in info sessions. Met with pilot participants. |
| Union President | Implementation Support/ Communicator Presented pilot at union meetings and some school sites. Communicated with membership. Organized at school sites. Met with individual teachers. | Facilitator/ Communicator Convened meetings and facilitated discussions. Communicated work with stakeholders. Reviewed feedback. Proposed and voted on policy changes. Implementation Support Led trainings. Participated in info sessions. Conducted observations. Administered surveys. Met with pilot participants. |
| Union reps / teachers | Implementation Support One rep presented at union meetings and communicated with membership. Watched demo lessons. Took on greater leadership role when the committee reconvened. Contributors Pilot participants attended trainings and meetings to ask questions and provide feedback. Completed surveys. Participated in evaluations. | Communicator One union representative took on some leadership tasks, including engaging with union members. Contributors <i>Committee members</i> attended meetings and participate in discussions. Reviewed feedback and proposed and voted on policy changes. <i>Pilot participants</i> attended trainings. Joined some evaluation committee meetings to ask questions and provide feedback. Completed surveys. Participated in evaluations. |
| School admin | Contributors Pilot participants attended trainings and meetings to ask questions and provide feedback. Completed surveys. Conducted teacher evaluations. Collaborated with the AS of HR. | Contributors <i>Committee members</i> continued to attend meetings and participate in discussions. Reviewed feedback and proposed and voted on policy changes. <i>Pilot participants</i> attended trainings. Joined some evaluation committee meetings to ask questions and provide feedback. Completed surveys. Conducted teacher evaluations. |

Facilitators and Communicators

The individuals who assumed leadership roles on the teacher evaluation committees in the preceding stages essentially maintained those positions. In Sumac, the Assistant Superintendent of HR and Union President continued to facilitate committee meetings. They both maintained their roles as “communicators,” with support from another union representative and district HR administrator. As discussed earlier, they engaged district stakeholders, both pilot participants and the larger community, by soliciting feedback and providing updates. One union representative described the communications tasks she engaged in:

I've tried to energize and mobilize and educate some key people at each school to make sure that word is out about their rights and privileges.

Union presidents in both districts communicated with their members throughout implementation, making sure that teachers were aware of how the new policies affected them.

When Juniper formed its third teacher evaluation committee, at the beginning of the second year of implementation, one union representative and the Assistant Superintendent of HR served as the co-chairs. The Assistant Superintendent of HR took the superintendent's place as the lead district representative, and the nominated him to act as the group's facilitator. The union representative was the only individual who served on all three of the district's committees and thus took on more leadership roles as the work progressed.

Implementation Leads and Support

As district reform efforts moved from the design stages to the pilot, individuals assumed new roles. In Juniper and Sumac, the Assistant Superintendents of HR served as the implementation leads, supported by other individuals, including the union presidents and key union representatives. In Juniper, the superintendent also provided direct implementation support, as did a district HR administrator in Sumac. The implementation leads served as the “point-persons” for the pilot. One individual described his role as follows:

I'm the point person for teacher evaluation. So I've been doing all the training. Again, working with both the principals and teachers, working with the union on it, doing those collaborative observations with principals, picking up evaluations, reading evaluations, talking about how evaluations are written. It's basically a lot of PD aspect of it, gathering input, write surveys, those think tanks.

In Juniper, the Assistant Superintendent of HR joined the district as they were finishing crafting the pilot and took a few months to get up to speed, but then quickly assumed this leadership role. Implementation leads took on many responsibilities, including running information sessions, conducting trainings and supporting principals. In Juniper, the implementation lead went to every school site to collect the evaluation forms and meet with administrators. The implementation leads conducted demo lessons and also demonstrated how to conduct the pre- and post-observation conferences.

The individuals who provided implementation support also participated in info sessions and trainings. They watched demo lessons in classrooms and had conversations with pilot participants. They may also have helped to prepare the pilot materials and administer surveys. In Juniper, the superintendent taught a demo lesson in a classroom and had the lead union representative evaluate him using the pilot rubric. Two individuals who provided implementation support in Sumac – the union representative and district HR administrator – had been involved in the teacher evaluation work from day one, and their continued involvement provided consistency. Because of that, they were also identified by others as leaders.

Contributors

In Sumac, teacher evaluation committee members were involved in year one of the pilot. The committees' work shifted from designing policy to quality improvement tasks. Committee members contributed by meeting throughout the year, supporting the information sessions and trainings, and listening to feedback from pilot participants. In this stage, pilot participants became contributors, as well. In Sumac, pilot participants contributed by attending committee meetings throughout the year; individual principals and union site representatives from pilot schools were invited by the committee. At the end of the year, all pilot participants in Sumac attended a meeting to discuss implementation. Similarly, in Juniper, pilot participants were invited to a meeting to provide feedback, though regular committee meetings were not being held.

Stakeholder Relationships

Prior to the implementation stages, labor-management relationships were improving in Sumac, but deteriorating in Juniper. As the teacher evaluation pilots were rolled out, more individuals became involved in reform efforts at this stage and school-level relationships became more relevant. Stakeholders from both districts noted that teacher-administrator relations varied by school site, and that those relationships impacted the implementation of reforms. In this section, I discuss how the relationships between the districts and teachers unions evolved in the first year of the teacher evaluation pilots and comment on their status as of the end of 2014.

Sumac

Throughout year one of the pilot, Sumac's teacher evaluation committee continued to meet. By this stage, many of the committee members had been working together for two years. As a result, their interactions had become collegial and they were able to have open dialogue. The success of the committee in fostering labor-management collaboration was also recognized by individuals who had not participated in the teacher evaluation reform efforts. The teacher evaluation committee had become an example of collaboration for the district as a whole. As one interviewee commented:

The work that that committee has done has gone great lengths in showing people around the district, our hardliners again in both camps, that it's much better to work together than to just try and battle stuff out at the table.

Their success in working together to change the teacher evaluation policy impacted the district's culture, by demonstrating that collaboration is both possible and beneficial to all stakeholders. The district began using this committee as a model when establishing new committees and the positive experience with the teacher evaluation work enabled dialogue in other policy areas.

After having worked on teacher evaluations together, the district and union were willing to take on other joint-initiatives. Sumac's commitment to fostering and sustaining labor-management collaboration was evident in their decision to hold a joint labor-management leadership summit in the summer before their teacher evaluation pilot began. The goal of the summit was to build trust and strengthen relationships beyond the context of the teacher evaluation work. As implied by the name of the summit, the district and the union planned it together. While the summit was not focused solely on teacher evaluations, it helped lay the foundation for successful implementation of the pilot, by getting all stakeholders on the same page. One participant said:

What was so wonderful about the leadership summit, we had teacher leaders from every school, and we had the principal and the assistant principals together in rooms talking about it and seeing what the future was going to bring.

The summit provided an opportunity for leaders from both sides to address perceptions and misconceptions held by the "other side." They tackled some difficult issues head-on in a panel discussion, as one individual described:

I'll give you an example. The teachers might say that the managers have their pets and they only want to do things for their pets, you know? And it was just real conversations. And then I said that, I hear managers saying that the teachers association's more, they're more concerned about maintaining employment for the people in their association as opposed to student learning. And people say that they really, really appreciate it because we started talking about things that usually you only say in your own group. And there's different perspectives on it, but both of us got to weigh in on both of those.

The fact that stakeholders from both labor and management were willing to have those conversations, and found the experience to be a productive and valuable, provides evidence of the trust and respect they were able to cultivate.

Still, relationships still were not consistent across all school sites. At some schools, teachers and administrators worked well together, while at others, the relationships were more tenuous. In the first year of the pilot, the district selected schools with exemplary labor-management relationships. District administrators who were interviewed recognized the importance of working with schools that have difficulties, as site-level interactions will influence continued implementation. In fact, the district received a grant to work on labor-management relationships at the school level. They intend to conduct trainings and field climate surveys at selected school

sites. This was a natural extension of the relationship-building work that took place in the teacher evaluation committee.

Sumac was proactive in engaging stakeholders throughout the first year of the pilot, getting feedback from a variety of individuals. Contract negotiations in the spring went quickly and smoothly and the district and union agreed to turn the pilot into a permanent system; the collaborative labor-management relationship enabled them to move forward. One individual noted that while bargaining could still be a difficult process, and they occasionally lost decorum, their interactions were never outright disrespectful. Leaders from both the district and union said that they have gotten better at solving problems collaboratively. However, they still experience challenges. Sometimes leaders may be hesitant to collaborate because they may be perceived by the individuals they represent as caving-in to the demands of the “other side.” One individual described that conflict:

So what happens in a relationship like that, because people have their suspicions and they have their perceptions of labor relationships with managers, they... I'm sure, [she] gets feedback from people, "You're just doing everything the managers want." And we might get feedback, "You're just doing everything the association wants," but if people were to look closely, they will see that we're doing right by kids. We try not to be positional, but we try to look at it: What are we doing to help students?

Occasional conflict should be expected, even when labor-management relationships are mostly collaborative. As observed in Sumac, the ability to resolve that conflict depends on the strength of those relationships.

As of the end of 2014, halfway through the second year of the pilot, district and union leaders were maintaining regular communication, making sure to share information and be transparent. This is part of a continued effort to maintain the positive state of labor-management interactions. Two union representatives were meeting with two HR administrators every two weeks, sometimes informally, and the superintendent meets with union and district leaders once a month. They still experienced bumps in the road, where some individuals might react too quickly, but they have learned how to address issues. While both sides do not always agree, they are willing to work together as professionals to address problems and resolve disagreement, as one interviewee said:

We don't always see it the same way, but that doesn't mean that the relationship is contentious at all. It just is, we just have professional differences and we try to work them out together. So I think that's qualitatively different than what it's been before.

Another individual echoed that sentiment:

We don't always see eye to eye on things, but we treat each other professionally and we're always willing to work together. That's the difference, and that's huge.

The positive relationship between the superintendent and union president, and between labor and management, more broadly, has not gone unnoticed. Sumac received recognition from an

external organization and were invited to present at a conference on labor-management collaboration, which provides further evidence of how much their relationship has improved.

Juniper

As discussed in previous stages, the labor-management relations in Juniper became strained after new union leadership was elected. In the summer before the first year of the teacher evaluation pilot, a third union president (UP #3) took office, and the climate did not improve. The current Assistant Superintendent of HR joined the district at the same time, so the two did not have a pre-existing working relationship, nor had either of them been previously involved in the teacher evaluation committees. In some ways, there was less collaboration over the course of the first pilot year, as the committee no longer held regular meetings. UP #3 did join the new Assistant Superintendent of HR in conducting information sessions at some school sites, but the Assistant Superintendent remained the primary leader of the pilot and presented to a few schools on his own. The union continued to express concern over consistency in implementing the new teacher evaluation across school administrators. In an attempt to build trust, the district invited the union to participate in the training they held for administrators. During that training, school administrators watched videos of teachers and rated their instruction on the rubric. Still, some stakeholders maintained that the evaluation was flawed and continued to question the district leadership's intentions. This made it difficult to move beyond a civil relationship to a truly collaborative one.

As in Sumac, teacher-administrator relations varied by school site, and there were some conflicts – both related to and unrelated to teacher evaluation – that affected implementation of the pilot in the first year. In fact, the number of permanent teachers participating in the pilot seemed to be associated with how active the union was in a particular school. At one school, tensions were especially high after the administration decided to dissolve a long-standing committee that included teachers in school-level decision-making. Teachers felt that their voices were no longer being heard, which contributed to their distrust of management. At the same school, there was a record-keeping snafu, where teachers were mistakenly informed that they would be evaluated, even though they were not due for an evaluation. This happened before the official information session about the pilot took place at their school. Because of that, permanent teachers may not have been aware that the pilot was optional for them and some may have felt pressured to participate. This mistake was perceived by some to be an intentional effort to coerce tenured teachers to participate in the pilot, as one individual commented:

According to my understanding, attempting to get as many people on this pilot as possible so that they could create the impression that people had chosen this pilot, because I think that they were trying to inflate the numbers of people being evaluated on the system. For whatever purpose, I don't know. But basically, they went in and they admitted to a huge error in the current standing of records keeping in the HR department and that they had straightened it all out.

This misunderstanding led to increased union mobilization against the reform, a behavior that aligns with contentious labor-management relationships.

While interactions were not always adversarial, labor and management were not perceived as being on the same page and distrust prevailed. Some district stakeholders assumed the union was trying to sabotage the pilot by discouraging participation, and some union stakeholders were convinced that the district was trying to coerce teachers to participate. Further, some union members remained skeptical of the district's original intentions in pursuing teacher evaluation reform. One individual described some ongoing concerns:

There's a faction of people who were distrustful that the process wasn't... They thought that the process was just looking for a way to fire teachers, to get rid of teachers, although it had no effect on ed code or tenure.

In addition, an increase in informal classroom visits by principals at some schools, which were not prescribed by the pilot evaluation, fed some teachers' existing distrust; they saw these drop-ins as another mode of evaluation. Despite district assurances, some teachers still feared that the evaluation process was going to be a "gotcha," or a way to penalize them. For these reasons, it was difficult to foster a positive and collaborative labor-management relationship.

Despite some stumbling blocks, there was also positive feedback from participants in Juniper's teacher evaluation pilot and the district and union agreed to formally adopt the new teacher evaluation policy in contract negotiations. They came to a tentative agreement in the spring of 2014, at the end of the first year of the pilot. Some individuals noted that negotiations, while lengthy, went more smoothly than in the previous year, which suggested that the labor-management relationship was improving. The negotiations team agreed to very specific guidelines about the structure of the third evaluation committee, which would be tasked with evaluating the second year of the pilot and making recommendations for improvements. Though some union leaders still had misgivings about the pilot, they found the structure of the committee agreeable, because it would be impossible for changes to pass without support from both district and union representatives.

The union in Juniper became less vocal about their concerns after the agreement was reached, making fewer comments at school board meetings and in other forums, as they did before implementation began. This suggests that they no longer saw the new system as "high stakes." Some observed the union becoming "hands-off," which is a behavior characteristic of civil, but not collaborative relationships. When the third teacher evaluation committee began meeting in the fall of 2014, there were signs that they would be more collaborative than the previous committee. However, what they are able to accomplish has yet to be seen. There were other signs that labor-management relations were improving. Contract negotiations in the fall of 2014 were completed much more quickly than in previous years, when they had carried on into the spring. They spent longer days at the bargaining table and did not need a mediator. Interactions were no longer contentious, as they had been in previous rounds of negotiations. Stakeholders on both sides acknowledged that drawn-out negotiations were not good for morale.

As one individual commented, the district and union were trying to improve their communication:

I think that at the end of last year, there were just some conversations that it can't go on like this anymore. So, I think transparency. I think... There's a desire to make the relationship work. And so, I think that that's been a part of it. Communication, you know.

Unlike in Sumac, where the teacher evaluation committee was a model of district-union collaboration, the impact of the teacher evaluation committee on labor-management relations in the district as a whole was less clear. While the most recent iteration of the committee was off to a good start, there was continuing conflict around other issues, including an outstanding labor complaint, which made collaboration challenging. Some stakeholders said that labor-management relations reached a low-point at the end of the first pilot year. One barrier to labor-management collaboration was the perception that the school board was particularly “anti-union.” Interactions between the union leadership and school board had been combative. Teachers expressed concern that the board wanted to erode their job security protections. One stakeholder said, “I’ve never seen anything like this before.” The relationship between district and union leadership remained tenuous, though not contentious.

Influential Factors

In the stages preceding implementation of the teacher evaluation pilots, the case study districts faced challenges posed by leadership turnover and contract negotiations. However, they did not rush the decision-making process and stakeholders were able to compromise and move forward with implementation. The decision to implement a pilot prior to adopting a permanent system was a successful approach in both districts. In the first year of the pilots, the districts received both positive and negative feedback from participants, and they responded by making quality improvements prior to entering contract negotiations once again. There were some unanticipated challenges with implementation. District culture continued to be relevant, but at this stage, school-level culture and politics emerged as important influential factors. There were some missteps with communication and coordination. Leaders’ efforts to communicate with and engage stakeholders were crucial to moving forward, as implementation could not be successful without stakeholder support. The districts faced different circumstances, but both were able to transition the pilot evaluation systems into permanent features of their CBAs during contract negotiations. In this section I discuss the key enablers of and barriers to progress experienced by the case study districts in early implementation.

Leadership and Relationships

Leaders – both of the district and the teachers union – set the tone for policy implementation and shaped how reforms were received by the district community. Further, their relationships,

and the degree to which they were willing to work together, affected the level of stakeholder support for their work and also the willingness of others to participate in pilot programs. Teachers union presidents and superintendents played crucial roles in shaping the labor-management relationship and setting the tone for reform efforts. As one stakeholder said:

A skilled superintendent know how to do that, knows how to bring the teachers union along on policy issues.

Sumac's joint-leadership summit in the summer before the pilot began signaled to the community that labor-management collaboration was a priority. It also gave leaders the opportunity to find ways to work together productively during the roll-out of the teacher evaluation pilot and other district initiatives. They maintained this collaborative approach throughout the year. Sumac's teacher evaluation committee continued to meet regularly and leaders met outside of the committee, as well. This approach demonstrated to the district community that the teacher evaluation reforms were a joint-effort and that the leadership on both sides were committed to its success. One individual described that evolution of labor-management relations in Sumac:

It has become over the recent years, a very positive relationship... And 90% to 95% of the time we have common interests, it's all about, "What are we doing for our kids? And how can we be the best that we can be?" And so, with the leadership that we have now in the district, and the school board that we have now, where we've been able to grow a relationship that is collaborative.

Stakeholders also recognized the importance of continuing to work on their relationships, keep meeting, even when they were busy or tempted to do other things. One stakeholder explained, "You have to keep the relationship going. That requires still an investment in time, constant investment in time." This can be difficult when there is turnover in leadership. In both Juniper and Sumac, the superintendents were relatively new to their districts when implementation began. As seen in the early reform stages, leadership turnover can threaten district-union relationships. Juniper experienced this challenge as new leaders joined the district and union, right as implementation was set to begin, and prior relationships were essentially reset. The change in leadership led to tension, as one individual said, "I think the tensions are at their highest point that they've been." Leaders have to be extraordinarily committed to change, especially when it is perceived as controversial, as they are ultimately responsible for motivating and mobilizing their membership. Juniper experienced problems when new union leadership, who had not been previously involved in the teacher evaluation work, was not enthusiastic about the pilot. Some stakeholders felt that the union was discouraging participation, which could be a barrier to the success of their reform efforts.

It was clear that district and union leaders did not see eye-to-eye and that colored their communication with the district community. Leaders in Juniper were less willing to issue joint-messages, which would signal a unified voice, as was done in Sumac. In Juniper, the union may have been wary of issuing joint statements and appearing to be in cahoots with the district

administration. One Juniper stakeholder described why it is important to keep that separation, saying, “The relationships don’t have to be adversarial, but I think there needs to be separation.” Another individual expressed a similar sentiment:

To me, it's as important as the balance of power in the Legislative, Judicial and Executive branch. Honestly, I really do see it that way and I think that when people misunderstand their role and begin to merge into this unified voice, it's very dangerous.

Given the characteristics of each district’s labor-management relationships during the first year of implementation, it is not surprising that Sumac’s contract negotiations went well, while Juniper’s did not. Juniper had to use a mediator to settle differences, and while the union eventually agreed to the new teacher evaluation pilot, they did not wholeheartedly endorse the new policy. One union representative described how the contract language adopted could be interpreted as adding a third evaluation option for permanent teachers, and did not see the new system as mandatory. This could be a problem moving forward.

Sumac, on the other hand, was able to move through negotiations quickly and smoothly. One stakeholder noted that this was a return on the investment they made during the previous reform stages, saying:

Because we had worked collaboratively ahead of time, because we had done all of that ahead of time, the shift was actually really simple. And the bargaining part was actually really simple, because we had done everything hand in hand.

Leaders in Juniper did recognize the damaging effects of their strained relationship, and between the first and second year of the pilot, tried to improve it, as described below:

I think that at the end of last year, there were just some conversations that it can't go on like this anymore. So, I think transparency. I think... There's a desire to make the relationship work. And so, I think that that's been a part of it. Communication, you know.

Leaders began to have more informal conversations and some individuals said that the relationship was getting better. As one individual said:

So, I think there's just been more communication. Our superintendent talks obviously to the union, has that line there, and I just think that it's become more positive. And the superintendent was on the negotiating team. He was the chief negotiator. It meant a lot. It was very unusual... And so I think that that really sent a strong message that it's important, and we want work together.

However, another stakeholder made a cautionary point. If leadership changes again, will the new policy be sustained? This is an ongoing threat to the stability of teacher evaluation and other reform initiatives in both school districts.

Communication and Coordination

Communication between leaders has been shown to be a crucial component of relationship-building. Communication between leaders and the broader district community is also essential

for effective district-wide implementation of a new policy. However, it can be imperfect and miscommunication can lead to challenges. The case study districts communicated the details of their teacher evaluation pilots through various means. Teacher evaluation committee members played an important role in communicating and coordinating the reform efforts. The individuals who assumed the roles of implementation leads and supports were also tasked with communication. In Sumac, the teacher evaluation committee kept meeting regularly throughout the first year of the pilot. This allowed them to address issues as they came up and prevent the dissemination of misinformation. In Juniper, the committee did not meet frequently in the first year of the pilot, which may have made communication more difficult. Implementation was supervised primarily by the one implementation lead, the Assistant Superintendent of HR. Some of the communication challenges in Juniper may have also been related to, or exacerbated by, the turnover in leadership, as previously discussed.

Stakeholders in both districts noted the importance of transparency surrounding the details of the teacher evaluation pilots and the work being done by the committees. One individual in Juniper reflected that the committee could have done more to be transparent:

In retrospect perhaps the committee itself should have been more transparent. We should have been talking more publicly or more openly about things that were being discussed. But I think it's a difficult thing to do because nobody has any time to write, or to go to a meeting to listen. And we tried to do a lot in a short period of time.

Sumac's teacher evaluation committee was able to communicate with stakeholders with a more unified voice, because they had a better working relationship and regularly released joint communications. One individual explained why joint-communication was important:

So there was no one person. And we did that intentionally because we didn't want teachers to feel like, "Oh, this is coming from the district." And we didn't want administrators to feel, "Oh, this is coming from the union." It's like, "No, this is coming from the evaluation committee."

Sumac applied that same philosophy to training. The district held trainings for teacher site representatives and administrators simultaneously, rather than providing information to the two groups in separate forums, as described here:

So before, once we selected those schools, then we brought in the principals and the union reps from all five of those schools to one meeting where, again, it wasn't, "We're sharing this with the principals and then we're going to go share it with the representatives." It's, "We're sharing everything with everyone at the same time." This is, we don't want there to be any hidden ideas. We don't want you to hear one message here and here a different message. And so, we shared the evaluation system with everybody, at one time, and gave them an overview of what it looked like.

This was an effective approach for Sumac. It ensured that participants from both labor and management received the same message about the purpose of the new teacher evaluation system, as well as the same information about implementation. Training, in addition to regular

communication and oversight, is important to ensure consistency, which emerged as a common concern, especially among teachers. Juniper took a different approach than Sumac did, first training administrators, and then offering multiple trainings for teachers. Juniper, a smaller district, opened the trainings to all teachers, not only the individuals who opted-in to the pilot. These trainings, for which teachers were paid, were in addition to the information sessions conducted by the Assistant Superintendent of HR – and sometimes the union president – at every school.

It is difficult to determine if one district's approach to stakeholder communication and training was superior, but consistent messaging proved to be important. Sumac's decision to communicate a message to all stakeholders in one forum was received positively by the pilot participants. Trainings set the tone for the pilot and are a way to get stakeholders on the same page. However, one training will not prevent misunderstandings from occurring, and communication with participants must continue throughout the pilot. There were some early challenges in implementing the teacher evaluation pilots resulting from miscommunication. Despite their efforts to engage the entire district community, stakeholders from both districts said that there were still individuals who were not aware of the pilot. Some stakeholders remained resistant to the policy changes in both districts. As one interviewee in Sumac said:

People who are not knowledgeable about it. People who maybe didn't read the email, haven't really been paying attention and they're just now discovering that they're going to be evaluated. Now they're pushing back because... And for whatever reason, I'm not judging them, but I think they're resistant to change because it's change. But I think once they go through it, they're fine. No one's come back and said, "We will not do this. I refuse."

Juniper experienced some additional implementation challenges related to communication, or miscommunication. They had difficulty recruiting permanent teachers to participate and recruitment was complicated by misinformation spreading at some school sites. Union representatives said that some teachers were under the impression that the pilot was mandatory because of messages they received from their principals. Miscommunication may have led to the snafu at one school in Juniper, where teachers who were not due for an evaluation were told to participate. Some teachers felt they were being tricked into participating by hiding the fact that the pilot was voluntary. To address the confusion, UP #3 went to individual classrooms of tenured teachers to clarify that the pilot was indeed optional. While not intentional, missteps at the pilot stage can compromise future success or generate backlash. One individual described how some teachers believed the district would use the new evaluation policy to get rid of teachers:

I mean I think that's the goal of any teacher evaluation system, and I don't think that it's limited to ineffective teachers. I think it's limited often to... I think sometimes the goal of an evaluation system is to be able to identify difficult teachers as ineffective in order to move them along.

Despite efforts to communicate with stakeholders, some may remain skeptical of the district's intentions. However, consistent communication between the teacher evaluation committee and the district community at every stage in the reform process may help provide clarity.

Approach: Pilot Before Permanent

It is not a given that a new policy will be piloted before being universally implemented; however, the two case study districts were wise in their decisions to take that approach. By rolling-out their new teacher evaluation systems incrementally, they had the opportunity to identify areas for improvement and cultivate stakeholder support. One stakeholder described the benefit of the pilot approach:

And so, that's why we were phasing it in strategically so that we can eliminate some of the barriers or things that we didn't anticipate and also be able to give greater information to the next phase and say, "Okay, this is what we learned, this is how you can do it more effectively."

Should changes be made after the first year of the pilot, the participants would have to be re-trained. This could be a logistical headache had districts skipped the pilot stage and moved directly to full implementation. Additionally, by beginning with a pilot, the “stakes” for teachers were kept low, which may have made some individuals more amenable to participating. The pilot was optional for permanent teachers in Juniper, and permanent teachers at pilot schools in Sumac could opt-out. In Sumac, there was the additional stipulation that if the teacher did not like the evaluation in the end, she did not have to accept the result.

Both districts had structured roll-outs with set timelines, but they took different approaches to selecting pilot participants. The individuals who participate in the pilot, both teachers and administrators, will ultimately determine its success. As discussed, Sumac strategically selected pilot schools. They chose schools where there were already good relationships between administrators and union representatives, as one stakeholder explained:

And we picked schools where we knew there was a good relationship between the reps and the principal... So we picked those schools on purpose, because we wanted it to be successful in the first year. We didn't want to put something out and then pick random schools where there might be problems and people are like, "Oh yeah, this really works, forget this." So we did that.

This approach worked well for Sumac. It helped to build positive “PR” regarding the pilot around the district before the new system was implemented at more schools in the second year. Additionally, a poor outcome in the first year could threaten the future of the policy during contract negotiations, as one individual described:

A coaching model can easily become a gotcha model and clearly they didn't want to go into negotiations with a school where it had gone off the tracks.

This was not a barrier in Juniper and Sumac, as both were able to adopt contract language that made the pilot permanent. They were also able to increase the number of participants in the second year on implementation.

Because Juniper is a smaller district, limiting the pilot to a subset of schools may not have been feasible. By taking the “opt-in” approach, Juniper increased the probability that the permanent teachers who participated would be open to change. It is also possible that the individuals who self-selected would have had experience with similar teacher evaluations either in their education training programs or in other districts. Thus, they may not have felt threatened by the new policy. There is also a risk to this approach – that not enough individuals will volunteer. Though there was some resistance, Juniper was ultimately pleased with the number of teachers who opted-in. Juniper’s approach may not have worked in a larger district, like Sumac. If too many teachers volunteered there and were located at schools across the district, it may have been difficult for district administrators to train and support all the participating school personnel.

An additional benefit of the “pilot before permanent” approach was the opportunity to continue to strengthen labor-management relationships, at the district and school levels. In Juniper, the pilot itself may have promoted collaboration at schools. The district received positive feedback from participating principals and teachers about working together:

And the ones that have participated, I think that there are collaborative conversations. That survey data that we received was very positive in terms of recommending... They would recommend it to other teachers, it's been collaborative. They feel like their principal learned something from the process, their principal provided targeted feedback in relationship to their goals, those types of things.

Stakeholders did provide a word of caution regarding the pilot approach. It was important that they did not have a pilot that would go on forever. One stakeholder provided this advice:

The main thing is not to have a long lasting pilot. I like that this is something that isn't a five-year pilot. It's a two-year pilot, and it goes in. I think the committee works; the evaluation committee works.

Setting a timeline for the pilot and the transition to a permanent system was an enabler of progress in both districts.

Engaging Stakeholders in Quality Improvement

The importance of communication in the context of policy implementation has already been discussed. In addition to communicating with stakeholders, the case study districts were also proactive in soliciting feedback during the first year of the pilot. In order for a pilot to be successful, districts must be committed to quality improvement and must engage stakeholders in that process. This is not only essential for making the policy more effective, but for ensuring that

stakeholders will be committed to implementing the policy once the pilot period is over. One individual noted the importance of securing broad stakeholder support:

But you realize that there are several stakeholders, if you will, that are part of the program and that if they don't believe to the level in which you believe, it can fall apart.

Juniper and Sumac solicited stakeholder feedback through a variety of channels, and received both accolades and constructive criticism. Sumac invited participants to committee meetings and made changes along the way. They were able to do that in the first year because the new evaluation system was not yet written into the CBA. Juniper also held meetings and conducted a survey of participants. In both districts, there were teachers who did not support the new system, which is why ongoing efforts to engage them in continuous quality improvement are essential. One individual commented on lingering concerns among teachers:

And although we've tried to problem solve those concerns, there is a percentage of our teachers that do not feel that it's a good thing for them. The way they had it before was much better. They were much better protected in the past than they would be in the future. And in short, they're right. They're absolutely right. It was better in the past for them. The future will be a little more volatile for them,

Implementation leads had to combat persistent misconceptions about the new policies. One stakeholder described having to repeatedly assure teachers that student achievement data would not be used to calculate their final rating:

I think a lot of teachers were nervous about it because they assumed that it was in there. It took us explaining, "No, it's not. We have lots of other best practices that are included in there, but student assessment is not a certain percentage of your evaluation."

Teachers may have been especially anxious about the potential role of student achievement data as they were simultaneously preparing for new standardized tests aligned with the Common Core standards. One stakeholder in Sumac relayed that concern:

The Common Core has made it harder, not easier. That's been a counter force actually in my opinion because teachers are feeling a little overwhelmed in that, not only is there a new accountability system, but now what is this whole new... I mean testing is new now and common core is new and nobody really understands and I mean we will roll it out, we'll train you. But, there is this feeling of, not only we are being evaluated in a new way but we are being asked to teach in a new way and that's not particularly good.

Efforts to solicit feedback must be accompanied by a response to stakeholder concerns. Unanticipated challenges are to be expected with the implementation of new policies. In fact, pilots are designed to reveal inadequacies. For example, Sumac did not anticipate the amount of time that would be required of administrators to complete the evaluation process for all of their teachers. They addressed this concern by decreasing the number of required classroom observations and teacher-administrator conferences. The challenges encountered in the first year

of the teacher evaluation pilots were not setbacks for the case study districts; rather, they were opportunities for improvement.

School Culture

While district culture continued to influence the teacher evaluation reform process, school culture, which is not uniform within a district, became an important factor during the implementation stages. In Juniper, the rate of permanent teacher participation in the pilot varied across school sites. District-wide, only 25% of permanent teachers opted-in during the first year of the pilot. Some stakeholders observed that schools with stronger union ties had lower participation rates, with no permanent teachers opting-in at some sites. The union presence and teacher-administrator relationships are different at each school site, and that could influence participation. Other characteristics of school-level politics may be relevant as well. Some schools may be more receptive to change, in general. One interviewee reflected on the challenge that posed and asked, “How do we implement change in a place that’s very resistant to change?” Unrelated events on school campuses may make implementation of any new policy more challenging, as seen with the events that transpired at some campuses in Juniper.

Several stakeholders commented that elementary school teachers seemed more receptive to the pilot than secondary teachers, and more permanent teachers at elementary schools opted-in than at high schools. High school teachers may be more resistant to change because they’re used to having autonomy within their departments. High schools also tend to be larger, with more teachers to evaluate than in elementary schools. One interviewee in Juniper described why roll-out was easier at the elementary schools:

I think that there wasn't the same problem with the elementary school with confusion about who was on a schedule or not. I think there was a better organization, a smaller faculty and it's easier for them to keep track of things and people themselves to know.

High school administrators in Sumac also expressed concern over the time required to evaluate their large teaching staffs. This problem was exacerbated at schools with fewer administrators to support the principal. One administrator said of the pilot:

So, getting it done with fidelity has actually been pretty taxing. Given what goes on with the high school campus. But, of all the things that I do, this is the most meaningful that I do. It's so... I'm not resentful of the time commitment, but it has been a real jump to hyper-speed compared to what we did before. It's been a real big time commitment.

Some interviewees also speculated that there were generational differences that affect reactions to new policies, and that younger teachers and principals may have been more open to change. The fact that Juniper is a high-performing district may also have contributed to a general resistance to change. One stakeholder commented on that perception:

The attitude is, "Why do we need to improve? Our test scores are high. Our achievement is high. There is no need to do that and you're kind of just being a pain by pushing that upon us."

Sumac mitigated the risk of school-level problems by taking a strategic approach to selecting participants. However, as they expand implementation district-wide, they might encounter unanticipated challenges.

Chapter Eleven: Conclusions and Recommendations

This dissertation presented three case studies of school districts in California working to reform their teacher evaluation policies, and explored the labor-management interactions throughout that process. My aim was to identify the roles played by various stakeholders over the course of these reform efforts and to understand how their individual and collective actions influenced the process and its policy outcomes. I proposed a conceptual framework to map the reform process and applied it to tell the story of each district's work. I identified a set of roles and relationship types, informed by my review of the literature and my case study data, and used those categories to analyze the actions of and interactions between stakeholders. Lastly, I discussed the factors that influenced the reform process and the labor-management relationships throughout, identifying both enablers of and barriers to progress and collaboration.

In this final section, I summarize the findings for each of my three research questions and present recommendations for policymakers, practitioners, and for further research. While there are limitations to the case study approach, and these the observations do not represent all districts in California, the lessons learned from the work of these three districts are relevant to others looking to pursue similar reforms.

Research Question 1

For each stage in the teacher evaluation reform process, I observed the roles played by key stakeholders, addressing the following research question:

- What roles do teachers unions and districts play in efforts to reform teacher evaluations and to what extent does the involvement of these groups shift throughout the reform process?

I identified leaders, participants, and observers and defined ten roles that were commonly observed in the three case study districts. Leaders took on the roles of initiators, conveners, facilitators, and implementation leads. Leaders were significantly or consistently involved during a given stage in the reform process. Participants acted as contributors, communicators, advisors, and implementation supports. An individual could be a contributor by serving on the teacher evaluation committee or by participating in the pilot program. The degree to which a participant was involved ranged from consistent to occasional. Observers, who were superficially involved in the reform process, acted as supporters or critics. Table 11.1 displays the roles I observed in the case study districts at the different stages in the reform process.

Table 11.1 Stakeholder Roles Observed by Stage

| | Level of Involvement | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|--------------------|---------------------------|-----------------------|--------------------------------------------|----------------------------------------------------|
| | | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Leader | Significant Consistent | Initiator Convener | Facilitator Communicator | Implementation Lead Facilitator Communicator |
| Participant | Consistent Occasional | Contributor | Contributor Advisor Communicator | Implementation Support Contributor Advisor |
| Observer | Superficial | Supporter | Supporter | Supporter Critic |

Leadership roles were typically taken on by superintendents, assistant superintendents of HR, union presidents, and union representatives. Johnson (1984) identified the individuals in leadership roles as the key players in school reform, whose actions and personalities dictate union behavior and labor-management relations. Table 11.2 displays the individuals serving as the primary leaders of the reform process in each district at the different stages.

Table 11.2 Individuals in Leadership Roles by District and Stage

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------|-----------------------------------------------------------------------|------------------------------------------------------------------|-----------------------------|
| | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Laurel | Superintendent AS of HR Union President Union Representative | AS of HR Union Representative | <i>n/a</i> |
| Juniper | Superintendent Union President ¹ | Superintendent Union President ² | AS of HR |
| Sumac | Superintendent Union Representative ¹ | AS of HR Union President Union Representative ² | AS of HR Union President |

Note: Juniper had multiple union presidents during the reform process and two held leadership roles. '1' denotes the individual who held that role at the beginning of the process and '2' denotes her successor. In Sumac, two different union representatives took on leadership roles. '1' denotes the representative who was involved in initiating the efforts and who served as union president in the subsequent stages. '2' denotes the other union representative who served as a leader of the teacher evaluation committee.

Other district administrators, union representatives/teachers, and school administrators participated in various ways. Participation grew during the implementation stage, as additional teachers and school administrators became involved with the pilot. Some school board members were observers, who were not directly involved, but supported the process. However, it was not clear if all school board members in the three districts were aware of the teacher evaluation reform efforts. Below, I describe the roles of the most important players – superintendents,

assistant superintendents of HR, union presidents, and union representatives – drawing comparisons across the case study districts and the different stages of the reform process.

Superintendents

Superintendents played different roles at different stages, sometimes serving as leaders and other times taking a backseat; however, they wielded influence throughout the reform process. Table 11.3 compares the roles that superintendents played in each of the three case study districts at the beginning of the reform process, during the decision-making stages, and during early implementation. If more than one individual in a district took on a leadership role in a given stage, then I refer to them as “co-leaders” of that stage.

Table 11.3 Superintendent Roles by District and Stage

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------|-------------------------------------------|--------------------------------------------------------|----------------------------------------------|
| | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Laurel | Co-Leader Initiator Convener | Participant Advisor | <i>n/a</i> |
| Juniper | Co-Leader Initiator | Leader/Co-Leader Facilitator Communicator | Participant Implementation Support |
| Sumac | Co-Leader Initiator Convener | Participant Advisor Communicator | Participant Advisor |

In all three districts, superintendents took on leadership roles in the beginning of the process, but their direct involvement waned as the reform process progressed toward implementation. They were instrumental in getting the reform efforts off the ground, playing the role of “initiator.” In the middle decision-making stages, the superintendents in Laurel and Sumac scaled back their direct involvement, though they remained engaged in the reform efforts, receiving updates from their representative, occasionally attending meetings and serving as advisors and/or communicators. On the other hand, the superintendent in Juniper continued to lead the process, acting as a facilitator of the teacher evaluation committee until implementation.

As indicated in the table, the superintendents were not the only “leaders” during any given stage. However, for a period of time, Juniper’s superintendent was the sole facilitator of the district’s teacher evaluation committee. When Juniper’s second teacher evaluation committee was formed, it was officially chaired by the superintendent and a union representative.

Assistant Superintendents of HR

Aside from the superintendent, the Assistant Superintendent of HR was the most involved of the district administrators. This is not surprising, given that personnel evaluations fall under the purview of Human Resources. Table 11.4 displays the roles played by those individuals.

Table 11.4 Assistant Superintendent of Human Resources Roles by District and Stage

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------|-----------------------------------------------|-----------------------------------------------|------------------------------------------------------------------------------------|
| | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Laurel | Co-Leader Convener | Co-Leader Facilitator | <i>n/a</i> |
| Juniper | | Participant¹ Contributor | Leader² Implementation Lead Facilitator* Communicator |
| Sumac | Participant¹ Contributor | Co-Leader^{1,2} Facilitator | Co-Leader² Implementation Lead Facilitator Communicator |

Note: Both Juniper and Sumac had two Assistant Superintendents of HR during the reform process. '1' denotes the individual who held that role at the beginning of the process and '2' denotes his successor.

*Juniper's second AS of HR took on the role of "facilitator" when the third teacher evaluation committee was formed as the pilot transitioned into the second year.

While the Assistant Superintendents of HR were not the initiators of the teacher evaluation reform efforts, they assumed greater responsibility as the work progressed. Moving into the goal-setting and other decision-making stages, Laurel and Sumac's Assistant Superintendents of HR served as co-facilitators of the teacher evaluation committees with their teachers union counterparts. In Juniper, where the superintendent facilitated the committee, the first Assistant Superintendent of HR participated as a committee member, rather than a leader. Juniper's second Assistant Superintendent of HR took on a facilitation role when the district's third evaluation committee convened in the second year of the pilot. In Juniper and Sumac, these individuals took on the lead role in implementing the teacher evaluation pilots. They also played an important role in communicating details of the pilot with the district community.

Teachers Union Leaders

Union leaders in all three districts were actively involved from the beginning; however the degree to which the union drove the reform process varied across districts. Leadership roles were filled by both union presidents and key union representatives. Union leaders played an important role in communicating the teacher evaluation work with union members and the district community as a whole. They also served as facilitators of the teacher evaluation committees. Other union representatives were involved as committee members or pilot participants. The roles played by union presidents and union representatives are presented in Table 11.5 and Table 11.6, respectively.

Table 11.5 Union President Roles by District and Stage

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------|-----------------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------------------------------|
| | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Laurel | Co-Leader Initiator Convener | | <i>n/a</i> |
| Juniper | Co-Leader¹ Convener | Co-Leader² Facilitator | Participant³ Implementation Support Communicator |
| Sumac | Participant¹ Contributor | Co-Leader² Facilitator Communicator | Co-Leader² Implementation Support Facilitator Communicator |

Note: Juniper and Sumac had multiple union presidents during the reform process. '1' denotes the individual who held that role at the beginning of the process and '2' denotes her successor, and in Juniper, '3' denotes the current union president.

Table 11.6 Union Representatives Roles by District and Stage

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------|----------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------------------|
| | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Laurel | Co-Leader Initiator Participants Contributors | Co-Leader Facilitator Communicator Participants Contributors | <i>n/a</i> |
| Juniper | Observers Supporters | Participants Contributors | Participants Implementation Support Contributors |
| Sumac | Co-Leader¹ Initiator Participants Contributors | Co-Leader² Facilitator Communicator Participants Contributors | Participants Communicator Contributors |

Note: In Sumac, two different union representatives took on leadership roles. '1' denotes the representative who was involved in initiating the efforts and '2' denotes the union representative who was elected as union president while already holding a leadership role in the teacher evaluation committee.

I observed different union responses to reform efforts in the case study districts that aligned with the response options proposed by McDonnell and Pascal (1988). The authors found that, in general, state and local unions generally accommodated reform proposals, rather than playing an active role in supporting or opposing new initiatives. However, the union leaders in all three case study districts were interested in shaping new policies, and took a proactive, rather than a reactive approach. The teachers union leaders in Laurel and Sumac were instrumental in getting the reform process started. They acted as “initiators,” in tandem with their districts’ superintendents. Laurel’s union leadership met over the summer during the “considering change” stage to discuss the role they wanted to play in shaping teacher evaluation policy. They then approached the superintendent about getting the reforms off the ground. In Sumac, the union was similarly interested in teacher evaluation policy and maintained a leadership role in every stage.

In Juniper, three different individuals held the role of union president during the teacher evaluation work, and the roles assumed by those leaders in the context of the reform process varied. The first union president (UP #1) helped convene the first committee, working with the superintendent. However, even under UP#1, the superintendent was considered the core leader of the committee. During later stages, new union presidents worked less collaboratively with the district leadership within the committee. While the second union president held the role of co-chair of the second committee, respondents viewed the superintendent as the lead, with the union representative taking a more oppositional stance. The union did not resist the reform efforts outright, as is sometimes observed in the literature, and their response turned out to be one of accommodation, as McDonnell and Pascal (1988) would expect. Union representatives in Juniper were involved in the negotiations that led to the formation of the committee, but were not actively involved in the first reform stage. Interviewee responses suggested that the teachers union did not feel a strong sense of ownership over the reform efforts, which could be due to their minimal involvement in considering change.

Conclusions

A range of stakeholders played a role in district teacher evaluation reform efforts and both the union and the district were represented on the committees tasked with guiding those reforms. The superintendent played an important role in initiating and supporting the process, and in keeping it moving forward. In the case study districts, stakeholders looked to see if an initiative was a district priority and if it had the superintendent's endorsement. However, in Juniper the superintendent was viewed as the sole leader of the initiative for a time, which proved to be problematic. In all districts, multiple individuals served as leaders, and their roles changed throughout the reform process. Reform leadership included representatives from the district and union. Union presidents and superintendents often delegated the committee leadership roles to others. Notably, the school board did not play a large role in driving the teacher evaluation reforms in any of the districts. However, the decision to not interfere may have been an implicit expression of their support for the work.

Research Question 2

My second research question addressed labor-management relationships:

- How do teachers unions and districts interact and how do their relationships change throughout the reform process?

In this section, I review the evolution of those relationships in each district over the course of reform process. I observed a full range in types of interactions, from collaborative to contentious, and the characteristics of those interactions changed over time. Table 11.7 categorizes the relationships in each district according to the typology defined in my conceptual framework (see Chapter 4) at the different stages in the reform process. I chose the category or categories that

best described the relationships during those time periods; however, it is possible that labor-management interactions deviated from these assignments at times. Relationships are not stagnant and “collaboration” is not a state of being. For example, a particular discussion could have been uncharacteristically contentious, when relations were otherwise collaborative. I also found variation in stakeholder perceptions regarding labor-management interactions. Not all stakeholders agreed on the state of the relationship at a given time. The categories I assigned reflected my analysis of all interview data and documentation, taking into account interactions within the teacher evaluation committees as well as in other forums, such as contract negotiations.

Table 11.7 Characteristics of Labor-Management Relationships Over Time

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------|----------------------|-----------------------------------------------|--------------------------|
| | <i>Stage 1</i> | <i>Stages 2-4</i> | <i>Stage 5</i> |
| Laurel | Collaborative | Collaborative | <i>n/a</i> |
| Juniper | Civil | Civil/Contentious | Contentious/Civil |
| Sumac | Civil | Civil/Collaborative | Collaborative |

Relationships at the beginning of the process were informed by history and shaped by the arrival of new superintendents. In Juniper and Sumac, those relationships were later affected by additional leadership turnover, in both the union and the district. As shown in Table 10.7, the labor-management relations in Laurel were collaborative and mostly stable during the time they engaged in the teacher evaluation reform process. While Juniper and Sumac both began in a less positive state, they followed different trajectories. The state of labor-management relations in Sumac went from civil to solidly collaborative. In Juniper, the relationship went from civil to fairly contentious, and then improved again toward the end of the first year of implementation. Walton and McKersie (1991) noted that, in general, relationships improve over time, but that leadership turnover may inhibit progress. This is consistent with what was observed in Juniper. However, the type of leadership turnover experienced by Laurel and Sumac did not inhibit relationship-building. While Sumac elected a new union president during the teacher evaluation reform process, that individual had already been serving on the evaluation committee since its inception. This provided continuity, unlike what occurred when new union leadership was elected in Juniper.

The interactions I observed aligned with the range of behaviors identified in the literature and the definitions proposed in my conceptual framework. Collaborative relationships were characterized mutual trust and friendly and cooperative interactions. When the relationships I observed in Laurel and Sumac were collaborative, unions were proactive in reform efforts and districts made a concerted effort to reach out the unions. Communication was regular and informal and both sides were willing to compromise. When relationships were civil, trust was

limited and the parties were courteous and somewhat cooperative. In other words, they tolerated each other, but compromise was not a priority. In Juniper, when the state of labor-management relations was civil, the district and union were able to come to agreement on a teacher evaluation pilot, though some union representatives were reluctant.

When relationships were contentious, as they were for a period of time after Juniper's second union president took office, trust was absent. Interactions were adversarial and formal. The presence of contentious labor-management interactions did not mean that all stakeholders were unwilling to work with the other side. However, as Johnson (1984) observed, it only takes one combative key player to effect a formal and adversarial relationship. When the state of labor-management relations reached a low in Juniper, the union was seen as obstructive and took action to mobilize its membership against the teacher evaluation initiative, while the district was viewed as somewhat authoritarian.

Co-leadership was a strategy used in all three districts, and likely contributed to collaborative behaviors observed in Laurel and Sumac. As previously discussed, in all case study districts, individuals from the district and the union took on leadership roles in the reform process. For the most part, the process was jointly led by the two sides. This collaborative approach to decision-making suggests that a district is moving from traditional forms of unionism to a more professional unionism (Kerchner and Koppich, 1993). Juniper was the exception, as the Assistant Superintendent of HR was the primary leader during the implementation stage, with support from the union. Juniper was also unique in that the superintendent stayed in a leadership role throughout the middle decision-making stages. It was during those stages that labor-management relationships became strained.

Research Question 3

My third research question, and the one that may be of particular interest to policymakers and practitioners, addressed the factors that shaped the reform process:

- What influences the roles, interactions, and relationships between teachers unions and districts during the reform process and the policy outcomes of their efforts?

I identified many enablers of and barriers to progress and examined the factors that facilitated or impeded labor-management collaboration at the various stages in the reform process. District and union leaders faced challenges at every stage, as they worked to resolve conflict, build relationships, and maintain momentum. In the first stage, stakeholders had to agree to address problems with the status quo and meaningfully consider change. After getting stakeholders on board with the initiative, they had to find common ground and plan their approach as they set goals in the second stage. As they moved to define and discuss options in the third stage, stakeholders had to navigate challenging conversations and find ways to overcome disagreements. Coming to agreement in the fourth stage relied on a mutual willingness to compromise, sometimes in the context of broader contract negotiations. In the fifth stage, when

implementation began, the key stakeholders had to be strategic in building support throughout the district and addressing unanticipated challenges and quality issues. Reaching this stage took a significant investment from and ongoing commitment of many individuals.

While all districts had some success at fostering labor-management collaboration, or at least cooperation, a collaborative relationship alone did not guarantee progress. In fact, as seen in Laurel, a positive relationship may have contributed to a desire to not “rock the boat,” which ended up stalling their work. Still, district efforts to build collaborative relationships often had the secondary effect of moving the reform process forward. In this section, I summarize ten lessons learned from my analysis of influential factors, related to the following five categories laid out in my conceptual framework (Chapter 4):

1. Approach to reform;
2. Leadership;
3. Reform features;
4. District characteristics and culture; and
5. Timing and external events.

These factors had varying degrees of influence and proved to be important in different ways as the reform process progressed. Table 11.8 indicates which of five influential factors, and related themes, were relevant at the different stages.

Table 11.8 Influential Factors by Stage

| | Setting the Stage | Approach to Reform / Hammering Out Details | Implementation Year 1 |
|---------------------------------------------|-------------------|-----------------------------------------------|-----------------------|
| | Stage 1 | Stages 2-4 | Stage 5 |
| Approach to Reform | X | X | X |
| Joint-approach | X | X | X |
| Careful planning | X | X | X |
| Stakeholder engagement | X | X | X |
| Investment in goal-setting | | X | |
| Pilot before permanent | | X | X |
| Leadership | X | X | X |
| Leaders ready for change | X | X | X |
| Turnover | X | X | |
| Role of superintendent | X | X | X |
| Reform Features | | X | X |
| Avoiding controversy | | X | |
| Consider implementation | | | X |
| District Characteristics and Culture | X | X | X |
| District culture | X | X | X |
| School culture | | | X |
| Timing and External Events | X | X | X |
| Timing | X | | X |
| External events | X | X | |

Most of the factors were important at every stage, albeit in different ways and to varying degrees. However, there were some exceptions. For example, while planning was important at all stages, the reform features that were discussed and later adopted did not have an influence until the teacher evaluation committees were convened. The culture at individual schools proved to be important during implementation, but not in the earlier stages. The degree to which each factor was influential will be addressed in the discussion that follows.

Approach to Reform

Several characteristics of the districts' approaches to reform influenced their success at every stage in the reform process. These include the decision to form joint-committees, the time spent in the different stages, the development of goals and timelines, the decision to enact a pilot before a permanent program, efforts to engage stakeholders, and the dedication to quality improvement, among other features.

A Joint-Approach Allowed Relationships to Grow

All three districts took the same approach of forming joint labor-management committees to work on their teacher evaluation policies. Though I have no counter-examples to draw comparisons, there appeared to be many merits to this strategy. Engaging stakeholders from both the district and union from the early, formative stages was important for the case study districts, as it has been noted to be in the literature (Donaldson and Papay, 2012; Goe, Holdheide and Miller, 2011; Poole, 1999). A joint, district-union approach marked a departure from a history of labor-management conflict and distrust. In Sumac, stakeholders noted that the collaborative work of their teacher evaluation committee proved to the community that it was possible for the district and union to work together. The case study districts' decisions to form committees outside of the context of contract negotiations gave teachers a voice in district policy decisions, made them less resistant to change (Behrstock-Sherratt et al., 2013; White et al., 2012), and set the tone for the subsequent stages. Further, this approach gave districts the opportunity to build trust and grow relationships (Minnici, 2014).

The joint-approach made stakeholders amenable to exploring reform in the first stage and also enabled them to come to an agreement in the fourth stage. The original idea to consider change did not come from district administrators alone, nor was it externally motivated. One union stakeholder described the benefits of an internally-motivated and joint-approach:

And I think that's real change. It's not something forced on us by some group from the outside. It's something that our members wanted; it's something that the district wanted. And when you get the actual professionals, and we are professionals, in the classroom, giving them the ability to work together to make a better environment to where we can do what we need to do for our students, for our children, which is our focus. Then you're going to get good things that happen.

At the foundation of a collaborative approach is joint-leadership. District leaders did not present committees with specific policy proposals, but delegated that work to the group and allowed union representatives to take on leadership roles. Having teachers drive the process was key. In Juniper, the superintendent was the primary leader in the beginning and middle stages, while the superintendents in Laurel and Sumac served in an advisory capacity once the committees convened. Though it was not clear why Juniper's superintendent made that decision – there may not have been another district administrator willing or able to lead the committee – having the superintendent serve as the leader may have given the impression that the union was not equally driving the process. Not surprisingly, there was more skepticism of the district's intentions in Juniper, when compared to Sumac and Laurel.

It was also important that committees had representatives from various stakeholder groups, as they would ultimately need to work together to implement the new evaluation system. The joint-approach was also relevant during implementation when districts needed teachers to participate in the pilot programs. In Juniper, where there was less district-union collaboration, some stakeholders felt that the union leaders were discouraging teacher participation, rather than endorsing the new evaluation system. In Sumac, district and union leaders issued joint-communications and made sure they were distributing the same message to the district community.

Careful Planning Moved the Process Forward

Joint decision-making takes time and is not necessarily efficient. The reform process can stall if there are no concrete timelines and next steps in place, as observed in Laurel. Careful planning was therefore essential for moving the process forward and accomplishing objectives. Since none of the case study districts were acting in response to legal requirements, they had to set milestones and timelines for themselves. Without structure, committees risked lingering in one stage and losing the momentum needed to progress to the next stage. This risk was especially present at the goal-setting and defining and discussing options stages. Laurel's committee, while collaborative, stalled in the goal-setting stage. Sumac was more successful at balancing a sense of urgency and the need to address important issues in that stage. Their "go slow to go fast" approach and investment in developing common goals paid dividends in later stages. Had they moved too quickly, prioritizing efficiency over relationship-building, they may have encountered additional challenges later in the process. This lesson is consistent with the conclusions made by Gates and Vesneske (2011) after observing contract negotiations. Sumac spent time building relationships and setting norms that would be needed when debating policy options. This early investment in relationship-building and planning allowed them to move more quickly in later stages. Further, their labor-management interactions set off on a positive trajectory at the second stage, while they languished in Juniper.

Planning was also important in the later reform stages. Both Juniper and Sumac were deliberate in their decisions to implement a pilot evaluation system before adopting a permanent

system. This lowered the stakes associated with teacher evaluations and enabled the districts and unions to come to an agreement. Both districts agreed on a two-year timeline for the transition and set a date for full implementation, which ensured that the process would not stall. Further, this approach allowed them to continue to build relationships as they monitored and evaluated the pilot. Knowing that the pilot would be reviewed and potentially improved prior to full implementation made the teachers union more amenable to the agreement, even though some individuals remained skeptical.

Implementation Relied on Communication and Stakeholder Engagement

Stakeholder engagement was a significant enabler of the reform process at every stage. Regular communication among leaders and between the teacher evaluation committees and the larger district communities laid the foundation for relationship-building and ultimately for successful implementation. On the other hand, a lack of communication and transparency served as a barrier to productive relationships and stakeholder buy-in. Sumac's committee made explicit efforts to solicit stakeholder input throughout the decision-making process, while committee members in Juniper reflected that they could have done more to make their work transparent.

As discussed, the pilot approach allowed districts to engage stakeholders outside of the committee in quality improvement efforts, address misconceptions, and build support. While stakeholder engagement was important in all stages, it became particularly crucial during implementation when more individuals became involved as pilot participants. Successful implementation relied on clear and consistent communication between the leaders of the reform efforts and the district community (Goe, Holdheide and Miller, 2011), as well as among district and union leaders (Eckert, 2011). Negative perceptions of district leaders and misconceptions about a new policy can compromise implementation and may have affected the number of permanent teachers who volunteered to participate in Juniper's pilot. Teachers were particularly concerned about the role of student achievement data and the consequences that could follow negative evaluations. To facilitate open channels of communication, Sumac's teacher evaluation committee continued to meet regularly during the first year of the pilot, issuing joint-communications with a unified voice. By keeping the committee involved at every step and regularly communicating their work, Sumac was able to maintain transparency. Both Juniper and Sumac provided trainings for pilot participants and organized forums to solicit feedback. There were some unanticipated challenges in the first year that participants brought to the districts' attention. This allowed for quality improvements that would increase the chance of successful implementation and stakeholder support in the future.

Leadership

The outcomes of any reform effort depend on people – from the individuals responsible for leading the effort to those that will be tasked with implementation (Johnson, 1984). Individual personalities and beliefs influence labor-management interactions (Walton and McKersie, 1991).

Leaders, of both the district and the teachers union, had the largest influence on the reform process of any stakeholder, and were influential at every stage. Their actions influenced whether or not the reforms got off the ground, the degree to which labor and management were willing to work together throughout the process, and whether or not various stakeholders supported the reform efforts.

Leaders Had to Be Ready for Change and Willing to Work Together

Leaders must be committed to the reform process and willing to collaborate (Donaldson, 2011; McLaughlin, 1990). They must make teacher evaluation a priority, even while faced with competing goals and other initiatives. Their attitudes influence how the broader community views the reform efforts and whether individuals will choose to participate in the pilot and endorse the proposed policy changes. Leaders can implicitly discourage participation and support if they do not fully buy-in to the new system, as was seen with the third union president in Juniper. In the case study districts, superintendents showed their support by attending committee meetings and relaying the work to the district community.

Labor-management collaboration in the case study districts began with leadership. District and union leaders had to be willing to work together, build relationships, and then continuously work to sustain them. Having a pre-existing relationship prior to engaging in new reform efforts was helpful, but not a necessity. Leaders recognized that they had to make relationship-building a priority even when they were busy. They did this by scheduling regular meetings with their counterparts and maintaining frequent informal dialogue. Sumac's leaders went a step further and organized a joint-leadership summit after recognizing that a strained labor-management climate was having a negative impact on the district community. Still, a positive and trusting relationship is not enough if leaders are unwilling to depart from the status quo (Koppich and Kerchner, 2000). Leaders in Juniper and Sumac were willing to negotiate changes to their teacher evaluation policies, which had not been touched for years. Had either party not been willing to expand the scope of contract negotiations, the teacher evaluation reforms would have never been implemented.

Turnover Created Opportunities for Change, Which Can Motivate or Impede Reform Efforts

A challenge faced by all three case study districts was leadership turnover, which occurred at the start of the reform process and in different stages throughout. This is an ever-present threat to the stability of district initiatives and the state of labor-management relationships. Koppich and Kerchner (2000) articulated the risks as follows:

Too often reform's tenure is dependent on personalities – the superintendent and union president have to be a good match. A change in leadership threatens whatever reform progress has been made. And this is particularly troublesome in urban districts where a superintendent can count on an average stay of no more than two or three years.

All three case study districts became engaged with changing their teacher evaluation policies as new superintendents were taking office. At the first reform stage, turnover was actually motivating. In later stages, turnover became challenging. Leadership in all three districts changed in the middle stages of the reform process. Sumac had two different Assistant Superintendents of HR and two union presidents. Juniper also had two different Assistant Superintendents of HR and three union presidents. The hiring of new assistant superintendents was less disruptive, because it did not change the districts' goals, which were set by the superintendents. However, new union presidents brought with them different visions for their roles and also different policy positions and priorities. In Juniper, the election of new union presidents posed several challenges. The composition of the teacher evaluation committee changed, and new members wanted to reevaluate previous agreements. Communication became more difficult and relationships had to be rebuilt. Since the final meeting of Laurel's teacher evaluation committee, two district administrators who were instrumental in driving the process have left the district. This may make it challenging to restart the reform efforts that have since stalled. Future turnover in any of the districts could be a setback to the progress they have made in improving labor-management relations (Walton and McKersie, 1991).

Long-term stability in leadership helps sustain collaborative relationships (Rubenstein and McCarthy, 2011). Still, while turnover is not something districts have full control over, they may be able to take advantage of transitions in leadership as opportunities for change. When they joined their new districts, all three superintendents put teacher evaluation policies on their agendas, enabling reform efforts to proceed. Since none of them had pre-existing relationships with their counterparts in the teachers unions, they had the opportunity improve the state of the labor-management interactions they inherited, though there was no guarantee that those relationships would be easy to foster.

Superintendents Had to Carefully Consider Their Roles

As discussed, all three superintendents played a role in initiating the teacher evaluation reform efforts. However, they chose to be involved in different ways once the committees began meeting during the middle decision-making stages. The superintendents in Laurel and Sumac stepped back and allowed committee members to take ownership of the work. Though they did not serve as leaders of the reform, they remained involved as advisors and showed their commitment. They listened to the committees' ideas without advocating for a particular policy option and that may have helped them gain trust. Having the superintendent serve as a leader of the committee may have proved problematic in Juniper. Though her commitment was evident, there was a missed opportunity to allow other stakeholders to take ownership of the reform efforts. Though it is difficult to say whether the superintendent's role resulted in some of the challenges experienced in Juniper, it was clear that they struggled with labor-management relations more than Sumac did.

Whether or not they served as leaders of the reform process, superintendents set the tone for how those efforts would proceed. Their involvement and support was crucial. Stakeholders looked to superintendents and assessed their commitment; they wanted to know that their work is supported and that it will be worthwhile. Individuals paid close attention the superintendent's communicated priorities and even the specific language they used. Negative perceptions of the superintendent's motives for pursuing or endorsing reform compromised stakeholder support.

Reform Features

Avoiding Controversy Facilitated Progress, But May Have Limited Scope of Reform

The reform features of the districts' teacher evaluation pilots most directly affected implementation. The language of the policies influenced stakeholder reactions and the requirements of the new system shaped the experience of the pilot participants. Both Juniper and Sumac received positive and negative feedback about the features of the new evaluation from participants. School administrators expressed concerns about the time and resources required to implement the pilots with fidelity. However, the range of reform features discussed, not only the specific ones that would later be adopted, influenced the work in the middle decision-making stages.

In the goal-setting and defining and discussing options stages, the teacher evaluation committees agreed to keep some controversial topics off the table, which helped them avoid conflict and may have enabled them to move forward. In general, there was a clear distaste for using student achievement data to calculate a teacher's rating and also for merit pay among the teachers union representatives. District administrators tried to focus the conversations on professional growth, to reassure teachers that a new evaluation system would not be used to punish teachers, as that was a common concern. Teachers may be more resistant to policy changes when the stakes are higher. In the case study districts, teachers were concerned about how the results of the new evaluations would be used; the stakes were unclear. Not all difficult topics were tabled, and avoiding a discussion around controversial issues, such as linking teacher evaluations to compensation decisions, did not assuage some individuals' fear that those reforms will be proposed in the future. However, by narrowing the scope of the debate, the committees freed up time work through other difficult reform features. One of the most challenging tasks was writing the teaching standards that would underlie the teacher evaluations. Neither Juniper, nor Sumac adopted another district's evaluation model, though they did incorporate pre-existing standards into the tools they designed. The most problematic areas were the standards related to teacher professionalism. Even though individuals had strong opinions about the standards, both districts were able to compromise by agreeing that teachers would not be evaluated on all standards in one evaluation cycle.

There are downsides to this approach. Taking topics off the table narrows the scope of policy options considered and may hinder efforts to explore other policies in the future. It is possible

that by excluding some features outright, the system will be less effective at achieving its desired outcomes, such as improving student achievement. Further, by choosing to avoid certain discussions, districts left the window open for policy debates and potential conflicts at later stages, which could compromise implementation. Limiting the scope of discussions may make some individuals feel that their voices are not being heard. One may argue that resolving disagreements is an important part of the process of designing effective policies. Learning to navigate conflict, which is inevitable, is an important skill that supports labor-management collaboration.

District Characteristics and Culture

Because of my small sample size, it was difficult to link certain district characteristics with specific outcomes. For example, one characteristic that I expected would have an impact on the reform process and labor-management relationships was district size. Johnson's (1984) observation about small districts being better able to collaborate was consistent with what I heard from several stakeholders in Laurel. They felt that their smaller size helped them build relationships. However, that did not seem to be the case with Juniper, an even smaller district. Additionally, Sumac, the largest district in my sample was successful at fostering positive relationships and making informal agreements; they were able to start the teacher evaluation pilot without a formal MOU, while Juniper had to settle that in contract negotiations before beginning. While small size can make it easier for stakeholders to interact and get to know each other, it does not guarantee collaboration. At the same time, large districts are not incapable of building collaborative labor-management relationships.

Still, it is important to consider the unique characteristics and circumstances in each district when proposing policy changes. Each district in my sample had a distinct history and culture that did influence stakeholder interactions at every stage, which I discuss below.

District Culture Affected Stakeholder Attitudes Toward Change

It is not uncommon for school districts to have had experienced labor-management conflict at some time in their history. Wounds from past conflict may have a lasting effect on stakeholders' willingness to trust each other and embrace change. Districts have to make a concerted effort to overcome the past, but as observed in the case study districts, that is not impossible. While the teacher evaluation committees operated outside of contract negotiations, a district's approach to collective bargaining affected labor-management interactions in other areas, particularly in the context of the teacher evaluation committees during the middle decision-making stages. Laurel and Sumac had adopted the IBB approach to negotiations, and that experience prepared them to collaborate on teacher evaluation reform. The IBB process served as a model for approaching reform. It encouraged stakeholders to be less positional and provided them with strategies for working through disagreements. As Rosenberg and Silva (2012) observed, teachers may have more positive feelings about their evaluations if they perceive the relationship between their

union and district to be collaborative and trusting. Sumac's long history with IBB aided them in structuring their approach to their teacher evaluation work. Laurel was in the process of transitioning to IBB when their committee began meeting, and some participants had recently completed related training. Laurel had also previously moved to a standards-based evaluation system, which may have prepared and motivated them to engage in a joint reform effort again. Juniper, on the other hand, followed the traditional model of industrial-style bargaining and struggled with collaboration during contract negotiations and within the committee. They may have benefitted from exposure to and success with more progressive types of bargaining.

District culture is tied to many district characteristics. Some attitudes toward reform observed in the case study district were related to student performance. It is harder to justify a need for change in a high-performing district than in a district struggling to raise student achievement. As a result, stakeholders in those districts may be less willing to accept reforms. Some stakeholders in Juniper and Laurel, the higher-performing districts, wondered why there was concern over ineffective teachers if student achievement was high. In Sumac, a lower-performing district, stakeholders were subject to external pressure and that may have brought them together; district and union stakeholders shared a common desire to raise student achievement and combat negative public perceptions. This contributed to their willingness to work together on teacher evaluation, helping them move forward from the first reform stage and later come to an agreement on a new policy.

Implementation Depended on Individuals and Schools

Implementation of teacher evaluations and many district policies ultimately happens on school campuses and depends on the actions of individual teachers and administrators. While not a barrier in the decision-making stages, school culture became an important influence once the case study districts began their pilots. Johnson (1984) observed that teacher-principal relationships influenced the way contract provisions are interpreted and enacted on individual school sites. Sumac expected implementation to go more smoothly at school sites where the teacher-principal relationships were already strong and positive, and strategically identified and recruited those schools for their pilot. Juniper, with fewer schools to choose from, opened the pilot to teachers at all schools. Stakeholders in Juniper noticed differences in the way the pilot was received at different campuses. Implementation went more smoothly at elementary schools than high schools. Participation was lower at campuses with a more active union presence. Some stakeholders also noticed generational differences, that younger teachers were more receptive to the new teacher evaluations, as they may have already been exposed to similar policies in their education programs.

While districts could not change the culture of schools to ensure the success of their pilots, they did make an effort to provide consistent training across all school sites and monitored implementation. Both districts wanted to make sure administrators were not interpreting requirements in different ways and that they were calibrated in their application of the teacher

observation rubric. However, one of the reasons for trying a pilot before implementing the new system district-wide was to identify problems that could be addressed in subsequent years. Therefore, some inconsistencies were expected. Teachers in all districts expressed concern about the consistency of implementation in their pre-existing systems and also in the pilot evaluations. This proved to be a greater issue in Juniper, where union leaders advocated for additional administrator training right before the pilot was set to begin. Discrepancies in the way the evaluation is implemented across schools can be perceived as unfair and compromise stakeholder support during the implementation stages. While those perceptions are problematic, the degree to which implementation was actually inconsistent across school sites was not yet clear. School-level conflicts in Juniper, some unrelated to teacher evaluations, may have fueled these concerns.

Timing and External Events

Timing Did Not Delay District Reform Efforts

Timing and external events had less impact on district reform efforts than the factors previously discussed. For the most part, district efforts were internally motivated, and that was seen as important. Timing can be fortuitous or debilitating, and can impact the reform process at any stage. However, it cannot always be controlled. It is difficult to anticipate leadership turnover or future state policies, so there is no perfect time to approach reform. In the three case study districts, the arrival of the new superintendents was timed well with the start of their teacher evaluation reform efforts. Districts did not time the work of their committees with contract negotiations, even though changes to teacher evaluation policies need to be written into district CBAs before they are permanent. Districts also did not delay their teacher evaluation reform efforts due to the confluence of other new policies being implemented, such as the Common Core standards. Nearly all interviewees did mention the Common Core standards, and while too many simultaneous reforms could overwhelm stakeholders, that did not seem to happen in the case study districts. That may be due, in part, to the fact that the districts chose not to include measures of student achievement in their new evaluation systems, as that would have been a more radical change. Still, it is possible that the timing of external events or state mandates could influence stakeholders' reactions to new policy proposals.

Stakeholders Paid Attention to External Controversies

Stakeholders in every district reported having knowledge of reforms to teacher evaluation policies in other districts across the country. Media coverage of reform efforts in other states and districts, especially controversial ones, captured stakeholders' attention. In some cases, peer districts served as examples of effective approaches to reform. Juniper's superintendent arranged for the teacher evaluation committee to visit a peer district that had success with revising their teacher evaluation system and a establishing collaborative labor-management relationship.

However, stakeholders were more concerned with districts that had faced challenges. Controversies publicized by the media contributed to skepticism about and resistance to teacher evaluation and related reforms. In particular, the *Los Angeles Times*' publication of individual teachers' value-added scores left many teachers in the case study districts uneasy and may have contributed to the decision to take student achievement measures off the table. Seeing the potential for adverse consequences raises the perceived stakes for teachers. External events have the potential for influencing the process at any of the reform stages. In the case study districts, concerns about external controversies were most relevant during the beginning stages, when policy options were being discussed.

Recommendations for Policymakers and Practitioners

“Reform” itself is not an inherently “good” thing and new policies can have good or bad outcomes. In this dissertation, I did not evaluate outcomes, nor did I intend to determine if teacher evaluation reform is “good” or “bad.” Instead, I focused on the process through which those reforms are developed and adopted and the effects labor-management relations have on that process. While I cannot provide recommendations on the types of reforms districts should pursue, my case studies did reveal many lessons about how reforms should be approached.

In the previous section, I explored the ways in which various factors contributed to successful reform efforts, and also how they made the process difficult. In the discussion below, I translate the lessons learned from my case study districts into recommendations for policymakers and practitioners. As the policy landscape surrounding teacher evaluations continues to change, it is likely that more districts will be interested in changing, or perhaps be forced to change their own teacher evaluation policies. State policymakers should keep in mind that the success of whatever teacher evaluation policies they choose to enact will depend on district and school-level stakeholders. As seen in my three case studies, district and school culture affects stakeholder willingness to change. Policymakers should pay attention to the controversies that compromised implementation in other states and try to avoid similar mistakes. They should be cognizant of the number of new policy changes districts must simultaneously comply with and implement, keeping in mind the limited resources available to them. Policymakers must also allow districts sufficient time to build support among their stakeholders and to adapt to new requirements, while providing incentives to keep the work moving forward. It would also be wise to give districts some autonomy in shaping the details of their teacher evaluation policies, as stakeholders may react poorly to externally-mandated systems that do not account for their unique local characteristics. Districts may be more amenable to complying with a policy requirement if they have some flexibility with implementation (Mead, Rotherham and Brown, 2012), and they may be more efficient at doing so if they are provided with resources and ongoing support.

State policymakers may also consider policies that explicitly encourage labor-management collaboration (Rubenstein and McCarthy, 2014), such as the provisions the ED included in the RTT competition, which required states to obtain union approval before submitting applications. Policymakers should also be cognizant of the impact that current efforts to weaken the power of public sector unions may have on teachers unions' reactions to reform efforts, regardless of whether or not those efforts are occurring in their own states. It could turn out, as Moe (2011) argues, that reducing the power of teachers unions will allow states to circumvent them and enact changes to teacher evaluations and other policies. However, as others have observed, such action may result in increased union organizing and obstructionist behavior, as they fight for their right to exist (e.g. Jacoby, 2010; Poole, 1999). Policymakers should consider whether it would be more productive for teachers unions to devote their resources to fighting for their right to exist, or to serving as partners in collaborative efforts to improve education policy.

At the district level, leaders should recognize that they play a crucial role in shaping the teacher evaluation reform process. They should take a joint-approach to designing new policies, engaging stakeholders at every stage. Union leaders can benefit from taking a proactive approach to change, and district leaders would be wise to work with their union counterparts from the beginning. Teachers will always be concerned about changes to the way in which they are evaluated and how the results will be used. If they are not involved in crafting those changes, they may be forced into a reactive position, opposing reforms (Donaldson and Papay, 2012; Poole, 1999). Some research contends that the presence of a union will make reform nearly impossible. However, as these case studies and others have shown, teachers unions can play important roles in driving reforms forward and fostering stakeholder support (Pogodzinski et al., 2015; Johnson et al, 2009; Bascia, 2005; Kerchner and Koppich, 1993). Districts should take advantage of that capacity. Leaders also need to show that they are committed to the reform process. Stakeholders can sense when one side is not enthusiastic or does not fully endorse an initiative. Districts will face many challenges and should be ready to address and resolve them. Below, I summarize my recommendations for the teacher evaluation reform process, many of which are applicable to reform efforts in other areas, as well.

Organizing and Managing Teacher Evaluation Committees

- Consider joint-leadership from the first stage and make sure the interest in reform is internally motivated and endorsed by both labor and management. Even if the effort is initiated by the district, it needs to have broad support. Select co-chairs and facilitators from both sides and ensure that the leadership role is jointly shared in practice, rather than in title alone.
- Plan ahead. Agree on a timeline, put it in writing, and work efficiently.
 - Be deliberate about committee structure, rather than allowing it to form haphazardly. Agree on who should be members and what constitutes consensus.

- Maintain a sense of urgency without rushing; the time required to build a successful working relationship should not be taken for granted. Slow, steady, and persistent wins the race.
- Invest time in goal-setting. Establish common language and agree on a definition of effective teaching.
- When considering policy options, always return to the agreed-upon goals. Keep the focus on student achievement and professional growth
- Be transparent. Make communication and stakeholder engagement a priority. Use various means for communicating with and soliciting input from the district community.
- Leadership turnover can be challenging. Keeping the committee composition consistent over time can help provide continuity.

Building Relationships

- A joint-approach can allow relationships to grow, but takes effort and commitment. Leaders must be ready for change and willing to work together.
- Collaboration is not a state of being that once achieved, maintains itself. Miscommunication and conflict will happen even when relationships are generally collaborative; resolve them quickly.
- Trust between the district and union is essential for collaboration, but is not enough for change. Leaders have to address tough issues, be flexible, and be genuinely willing to depart from the status quo.
- Avoiding controversy can facilitate progress, but may limit the scope of reform and delay disagreements that will eventually need to be resolved.

Implementation

- Start with a pilot, but have a clear pathway from pilot to permanent policy. Be careful not to get stuck in a cycle of pilots.
- Be aware of the time and resource requirements for school administrators to implement the process with fidelity.
- Use the pilot to build stakeholder support. Perhaps be strategic in selecting the first group of participants.
- Build a quality improvement process into the pilot. Refine the process before district-wide implementation. Be aware that a continuous improvement cycle will take significant time and resources.
- Develop a plan for communication and training. Make sure all stakeholder groups are receiving the same information at the same time, and that they have abundant opportunities to provide feedback.

- Maintain a focus on relationship-building. District and union leaders should be in close communication. The teacher evaluation committee should continue to meet regularly.

Recommendations for Further Research

While teacher evaluation is currently the subject of numerous research studies, less attention has been paid to the topic of labor-management interactions – in the context of teacher evaluation and other reforms. This dissertation bridged those two topics, but further research into the process through which districts and unions develop and implement reforms would be beneficial. I did not observe full implementation or the outcomes of new teacher evaluation policies. Policymakers and practitioners would benefit from further examination of the reform process and especially policy implementation, which has a bearing on outcomes. Researchers can observe implementation and explore the effects it has on stakeholder behavior and on the original policy objectives, namely student learning. Future studies can look for connections between a district's approach to the reform process and the impact of the policy that was implemented, thus connecting the early stages to the eventual policy outcomes. Researchers can ask if certain characteristics of the reform process influence the success of implementation, whether or not the policy is sustained, and the degree to which it achieves its goals. It would also be interesting to note if districts that successfully implemented new and effective teacher evaluation policies were then motivated to engage in other challenging reform initiatives, and whether or not they took a similar approach to new reform efforts.

District stakeholders would benefit from continued research that identifies effective teacher evaluation policies, and also effective strategies for policy design and implementation. I proposed several recommendations related to how districts should plan and structure their reform efforts, but my conclusions are limited to three case studies. Future research could apply a similar process-based conceptual framework to compare the reform process in other districts, which could support or challenge my conclusions. Using the stages of the reform process as a unit of analysis is a useful approach for generating specific and actionable recommendations for districts and for providing guidance on what they should anticipate and consider at each stage. This framework could then be applied to analyze the process through which districts approach other policy reforms. Additional case studies could be used to modify this framework, making it a more useful tool.

While states may debate whether or not they should change collective bargaining laws, local districts would benefit more from guidance on building collaborative and effective labor-management relationships. Even if future policies weaken unions' bargaining rights or political influence, teachers, and the associations that represent them, will remain important and influential stakeholders. Ultimately, the success of a district policy depends on the ability and willingness of individual teachers and administrators to implement it in their schools. I argue

that, instead of focusing efforts on debating whether or not teachers unions should exist, research should look at how unions can be engaged in and play a productive role in policy design and implementation. This question may be more immediately policy relevant to school districts struggling to comply with state requirements and improve student achievement.

While there continues to be research on the content of collective bargaining agreements and approaches to contract negotiations, significant labor-management interactions also occur away from the bargaining table. A closer look at how districts form and operate joint labor-management committees is important, as much district policymaking occurs in that context. Future research can continue to explore the characteristics of labor-management relationships and types of interactions in various settings, to help both sides work effectively toward the common goal of improving the quality of education provided to their students. While my study concluded while two districts were still in the teacher evaluation pilot stage, it would be helpful to understand how those relationships evolve – at the district and school-level – as implementation continues. Additionally, it may be helpful to observe decision-making within teachers unions, to better understand how they react to district reform proposals and how willing they may be to adopt more progressive and collaborative forms of unionism.

While my research focused on California, future studies could compare the roles of teachers unions in states with different types of collective bargaining laws and attitudes toward teachers unions. It would be interesting to examine how districts respond to state mandates regarding teacher evaluation policies and then compare that response to districts where reform efforts are internally-motivated, as they were in my case study districts. Researchers can also compare how districts in one state respond to the same mandate and examine the roles of the state and local teachers unions. A challenge to research in this area is that federal and state policies surrounding teacher evaluations and other personnel issues will likely continue to change and remain controversial. Further, policies related to both teacher evaluations and teachers unions will be shaped by the outcomes of pending lawsuits. Though a changing policy context might make it difficult to evaluate policy outcomes, it makes for an interesting setting to study labor-management relationships, which is worthy of further research.

References

- Aldeman, Chad, and Carolyn Chuong, "Teacher evaluations in an era of rapid change: From “unsatisfactory” to “needs improvement.”,” Bellwether Education Partners. Retrieved from http://bellwethereducation.org/sites/default/files/Bellwether_TeacherEval_Final_Web.pdf, 2014.
- American Federation of Teachers (AFT), "History," 2015. As of August 3: <http://www.aft.org/about/history>
- American Statistical Association, *ASA statement on using value-added models for educational assessment*, 2014. http://www.amstat.org/policy/pdfs/ASA_VAM_Statement.pdf
- Ballou, Dale, and Matthew G. Springer, "Using Student Test Scores to Measure Teacher Performance: Some Problems in the Design and Implementation of Evaluation Systems," *Educational Researcher*, Vol. 44, No. 2, March 1, 2015, 2015, pp. 77-86. <http://edr.sagepub.com/content/44/2/77.abstract>
- Bascia, N., "Triage or tapestry? Teacher unions’ work in an era of systemic reform," *International handbook of educational policy*, 2005, pp. 593-609.
- Bascia, Nina, *Unions in Teachers' Professional Lives: Social, Practical, and Intellectual Concerns*, New York: Teachers College Press, 1994.
- Behrstock-Sherratt, Ellen, Allison Rizzolo, Sabrina W Laine, and Will Friedman, *Everyone at the table: Engaging teachers in evaluation reform*: John Wiley & Sons, 2013.
- Blume, Howard, and Stephanie Cesar, "California teacher tenure and seniority system is struck down," *Los Angeles Times*, June 10, 2014. As of July 31, 2015: <http://www.latimes.com/local/education/la-me-teacher-lawsuit-20140611-story.html#page=1>
- Briggs, Derek, and Ben Domingue, *Due Diligence and the Evaluation of Teachers: A Review of the Value-Added Analysis Underlying the Effectiveness Rankings of Los Angeles Unified School District Teachers by the Los Angeles Times*, Boulder, Colorado: National Education Policy Center, 2011. As of July 30, 2015: <http://nepc.colorado.edu/publication/due-diligence>
- Brown, Emma, and Moriah Balingit, "Virginia pushed into debate of teacher privacy vs. transparency for parents," *Washington Post*, March 16, 2015. As of April 2, 2015: <http://wapo.st/1GJ53rY>
- Burnett, A., E. Cushing, and L. Bivona, "Uses of Multiple Measures for Performance-Based Compensation," 2012.
- Burroughs, Nathan, "Arguments and Evidence: The Debate Over Collective Bargaining’s Role in Public Education," *Center for Evaluation and Education Policy*, Vol. 6, No. 8, 2008.
- California Commission on Teacher Credentialing (CCTC), *California Standards for the Teaching Profession (CSTP) 2009*, Sacramento, CA, October, 2009. <http://www.ctc.ca.gov/educator-prep/standards/CSTP-2009.pdf>

- California Department of Education (CDE), "Student & School Data Files," 2014. As of December 9: <http://www.cde.ca.gov/ds/sd/sd/index.asp>
- , "Local Control Funding Formula Overview," 2015. As of August 24: <http://www.cde.ca.gov/fg/aa/lc/lcffoverview.asp>
- Carini, Robert, Brian Powell, and Lala Carr Steelman, "Do Teacher Unions Hinder Educational Performance?: Lessons Learned from State SAT and ACT Scores," *Harvard Educational Review*, Vol. 70, No. 4, 2000, pp. 437-467.
<http://hepgjournals.org/doi/abs/10.17763/haer.70.4.w17t1201442683k6>
- Center on Great Teachers and Leaders (CGTL), "Databases on State Teacher and Principal Evaluation Policies," 2013. As of August 9, 2013: <http://resource.tqsource.org/stateevaldb/>
- Chaplin, Duncan, Brian Gill, Allison Thompkins, and Hannah Miller, "Professional Practice, Student Surveys, and Value-Added: Multiple Measures of Teacher Effectiveness in the Pittsburgh Public Schools. REL 2014-024," *Regional Educational Laboratory Mid-Atlantic*, 2014. <http://files.eric.ed.gov/fulltext/ED545232.pdf>
- Cohen-Vogel, Lora, Li Feng, and La'Tara Osborne-Lampkin, "Seniority Provisions in Collective Bargaining Agreements and the "Teacher Quality Gap"," *Educational Evaluation and Policy Analysis*, Vol. 35, No. 3, September 1, 2013, 2013, pp. 324-343.
<http://epa.sagepub.com/content/35/3/324.abstract>
- Cooke, William N., "Improving Productivity and Quality Through Collaboration," *Industrial Relations: A Journal of Economy and Society*, Vol. 28, No. 2, 1989, pp. 299-319.
<http://dx.doi.org/10.1111/j.1468-232X.1989.tb00870.x>
- Cooper, Bruce S., and John Sureau, "Teacher Unions and the Politics of Fear in Labor Relations," *Educational Policy*, Vol. 22, No. 1, January 1, 2008, 2008, pp. 86-105.
<http://epx.sagepub.com/content/22/1/86.abstract>
- Cowen, Joshua M, and Katharine O Strunk, "The impact of teachers' unions on educational outcomes: What we know and what we need to learn," *Economics of Education Review*, 2015.
- Darling-Hammond, L., A. Amrein-Beardsley, E. Haertel, and J. Rothstein, "Evaluating teacher evaluation," *Phi Delta Kappan*, Vol. 93, No. 6, 2012, pp. 8-15.
- Darling-Hammond, L., A. Amrein-Beardsley, E.H. Haertel, and J. Rothstein, "Getting Teacher Evaluation Right: A Background Paper for Policy Makers," paper presented at Getting Teacher Evaluation Right: A Challenge for Policy Makers, Washington, D.C., American Educational Research Association and National Academy of Education, September 14, 2011.
- Darling-Hammond, Linda, and Frank Adamson, *Beyond the bubble test: How performance assessments support 21st century learning*, San Francisco: Jossey-Bass, 2014.
- Deery, S.J., and R.D. Iverson, "Labor-Management Cooperation: Antecedents and Impact on Organizational Performance," *Industrial & Labor Relations Review*, Vol. 58, 2004, p. 588.
- Donaldson, Morgaen L, "Principals' Approaches to Developing Teacher Quality: Constraints and Opportunities in Hiring, Assigning, Evaluating, and Developing Teachers," *Center for American Progress*, 2011.

- Donaldson, Morgaen L, and John P Papay, *Reforming Teacher Evaluation: One District's Story*, Washington, DC: Center for American Progress, 2012.
- Eberts, Randall W., and Joe A. Stone, "Teacher Unions and the Productivity of Public Schools," *Industrial & Labor Relations Review*, Vol. 40, No. 3, April 1, 1987, 1987, pp. 354-363.
<http://ilr.sagepub.com/content/40/3/354.abstract>
- Eckert, J. (Ed.), *Local labor management relationships as a vehicle to advance reform: Findings from the US Department of Education's labor management conference*, Washington, DC: US Department of Education, 2011.
- Fisher, Roger, William Ury, and Bruce Patton, *Getting to yes*: Simon & Schuster Sound Ideas, 1991.
- Fowles, Jacob, and Joshua Cowen, "In the Union Now: Understanding Public Sector Union Membership," *Administration & Society*, Vol. 47, No. 5, July 1, 2015, 2015, pp. 574-595.
<http://aas.sagepub.com/content/47/5/574.abstract>
- Goe, Laura, *Evaluating teaching with multiple measures*, Washington, DC: American Federation of Teachers, 2010. As of August 20, 2015:
<http://www.lauragoe.com/LauraGoe/EvalTchgWithMultipleMeasures.pdf>
- Goe, Laura, Courtney Bell, and Olivia Little, *Approaches to Evaluating Teacher Effectiveness: A Research Synthesis*, Washington, DC: National Comprehensive Center for Teacher Quality, 2008. As of August 20, 2015:
<http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf>
- Goe, Laura, Lynn Holdheide, and Tricia Miller, "A Practical Guide to Designing Comprehensive Teacher Evaluation Systems: A Tool to Assist in the Development of Teacher Evaluation Systems," *National Comprehensive Center for Teacher Quality*, 2011.
- Goldhaber, Dan, "Are Teachers Unions Good for Students?," in Rotherham, Jane Hannaway and Andrew, ed., *In Collective Bargaining in Education: Negotiating Change in Today's Schools*, Cambridge, MA: Harvard Education Press, 2006.
- Goldring, Ellen, Jason A. Grissom, Mollie Rubin, Christine M. Neumerski, Marisa Cannata, Timothy Drake, and Patrick Schuermann, "Make Room Value Added: Principals' Human Capital Decisions and the Emergence of Teacher Observation Data," *Educational Researcher*, Vol. 44, No. 2, March 1, 2015, 2015, pp. 96-104.
- Goldstein, Jennifer, "Easy to dance to: Solving the problems of teacher evaluation with peer assistance and review," *American Journal of Education*, Vol. 113, No. 3, 2007, pp. 479-508.
- Graham, Matthew, Anthony Milanowski, and Jackson Miller, *Measuring and Promoting Inter-Rater Agreement of Teacher and Principal Performance Ratings*, Washington, DC: Center for Educator Compensation Reform, 2012. As of August 20, 2015:
http://cecr.ed.gov/pdfs/Inter_Rater.pdf
- Grossman, P., S. Loeb, J. Cohen, K. Hammerness, J. Wyckoff, D. Boyd, and H. Lankford, *Measure for Measure: The relationship between measures of instructional practice in middle school English Language Arts and teachers' value-added scores*, National Bureau of Economic Research, 2010.

- Hamilton, L., B. Stecher, and K. Yuan, *Standards-based reform in the United States: History, research, and future directions*, DTIC Document, 2008.
- Hess, Frederick M, and Coby Loup, "Leadership Limbo," *New York: Thomas B. Fordham*, 2008.
- Hill, P.T., "The costs of collective bargaining agreements and related district policies," in Hannaway, J. and A.J. Rotherham, eds., *Collective Bargaining in Education*, Cambridge, MA: Harvard Education Press, 2006, pp. 89-109.
- Hoxby, Caroline M., "How Teachers Unions Effect Education Production," *Quarterly Journal of Economics*, Vol. 111, 1996, pp. 671-718.
- Humphrey, D. C., J. E. Koppich, J. A. Bland, and K. R. Bosetti, *Peer Review: Getting Serious About Teacher Evaluation*, Menlo Park, CA: SRI International, 2011.
- Jacoby, D.F., and K. Nitta, "The Bellevue Teachers Strike and Its Implications for the Future of Postindustrial Reform Unionism," *Educational Policy*, Vol. 26, No. 4, 2012, pp. 533-563.
- Jacoby, Dan, "Teacher Unionization in School Governance," *Educational Policy*, Vol. 25, No. 5, September 1, 2011, 2011, pp. 762-783. <http://epx.sagepub.com/content/25/5/762.abstract>
- Jester, Erin, "State makes teachers' evaluation scores public," *Gainesville Sun*, February 24, 2014. As of March 6, 2014: <http://www.gainesville.com/article/20140224/ARTICLES/140229783?Title=State-makes-teachers-8217-evaluation-scores-public>
- Jiang, Jennie Y., Susan E. Spote, and Stuart Luppescu, "Teacher Perspectives on Evaluation Reform: Chicago's REACH Students," *Educational Researcher*, Vol. 44, No. 2, March 1, 2015, 2015, pp. 105-116. <http://edr.sagepub.com/content/44/2/105.abstract>
- Johnson, S.M., M.L. Donaldson, M.S. Munger, J.P. Papay, and E.K. Qazilbash, "Leading the local: Teachers union presidents chart their own course," *Peabody Journal of Education*, Vol. 84, No. 3, 2009, pp. 374-393.
- Johnson, Susan Moore, *Teachers' Unions in Schools*: Temple University Press, 1984.
- , "Will VAMS Reinforce the Walls of the Egg-Crate School?," *Educational Researcher*, Vol. 44, No. 2, March 1, 2015, 2015, pp. 117-126. <http://edr.sagepub.com/content/44/2/117.abstract>
- Johnson, Susan Moore, and Susan M Kardos, "Reform bargaining and its promise for school improvement," *Conflicting missions? Teachers unions and educational reform*, 2000, pp. 7-46.
- Joseph, Nithya, "Teacher Trendline: Teacher Evaluations," *NCTQ Teacher Quality Bulletin Newsletter*, 2015. As of July 29, 2015: <http://www.nctq.org/commentary/article.do?id=153>
- Kane, T.J., and D.O. Staiger, *Gathering feedback for teachers: Combining high-quality observations with student surveys and achievement gains*, Seattle, WA: Bill and Melinda Gates Foundation, 2012. As of August 20, 2015: http://www.metproject.org/downloads/MET_Gathering_Feedback_Practitioner_Brief.pdf
- Kearney, Richard C., and Steven W. Hays, "Labor-Management Relations and Participative Decision Making: Toward a New Paradigm," *Public Administration Review*, Vol. 54, No. 1, 1994, pp. 44-51. <http://www.jstor.org/stable/976497>

- Kerchner, Charles, and Julia Koppich, *Teachers unions and educational reform*, Conflicting Missions, Organizing Around Quality: Brookings Institution Press, 2000a.
- Kerchner, Charles Taylor, and Julia E Koppich, *A Union of Professionals: Labor Relations and Educational Reform*, New York, N.Y.: Teachers College Press, 1993.
- , "Organizing around Quality: The Frontiers of Teacher Unionism," in Loveless, Tom, ed., *Conflicting Missions: Teachers Unions and Educational Reform*, Washington, D.C.: The Brookings Institution, 2000b, pp. 281-315.
- Kerchner, Charles Taylor, Julia E Koppich, and Joseph G Weeres, *United Mind Workers: Unions and Teaching in the Knowledge Society*, San Francisco, CA: Jossey-Bass Inc., 1997.
- Koppich, J.E., and C. Esch, "Grabbing the Brass Ring Who Shapes Teacher Policy?," *Educational Policy*, Vol. 26, No. 1, 2012, pp. 79-95.
- Koppich, Julia, "Addressing teacher quality through induction, professional compensation, and evaluation: The effects on labor-management relations," *Educational Policy*, Vol. 19, No. 1, 2005, pp. 90-111.
- Koppich, Julia E, and Charles Taylor Kerchner, "Rethinking Labor-Management Relations: It's a Matter of Trust, or Is It?," paper presented at Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, 2000.
- Koski, William S., and Eileen L. Horng, "Facilitating the Teacher Quality Gap? Collective Bargaining Agreements, Teacher Hiring and Transfer Rules, and Teacher Assignment Among Schools in California," *Education Finance and Policy*, Vol. 2, No. 3, 2007/07/01, 2007, pp. 262-300. As of 2015/08/04: <http://dx.doi.org/10.1162/edfp.2007.2.3.262>
- Lincoln, Yvonna S, and Egon G Guba, "Establishing trustworthiness," *Naturalistic inquiry*, 1985, pp. 289-331.
- Lockwood, J. R., Daniel F. McCaffrey, Laura S. Hamilton, Brian Stecher, Vi-Nhuan Le, and Jose Felipe Martinez, "The Sensitivity of Value-Added Teacher Effect Estimates to Different Mathematics Achievement Measures," *Journal of Educational Measurement*, Vol. 44, No. 1, Mar, 2007, pp. 47-67.
- Malen, B, JK Rice, C Jackson, K Hoyer, L Hyde, L Bivona, and J Sutter, *Implementation, payouts, and perceived effects: A formative analysis of Financial Incentive Rewards for Supervisors and Teachers (FIRST)*, Prince George's County, MD: Prince George's County Public School System, 2011.
- Mausethagen, Sølvi, and Lise Granlund, "Contested discourses of teacher professionalism: current tensions between education policy and teachers' union," *Journal of Education Policy*, Vol. 27, No. 6, 2012/11/01, 2012, pp. 815-833. As of 2015/08/04: <http://dx.doi.org/10.1080/02680939.2012.672656>
- Max, Jeffrey, Jill Constantine, Alison Wellington, Kristin Hallgren, Steven Glazerman, Hanley Chiang, and Cecilia Speroni, *Evaluation of the Teacher Incentive Fund: Implementation and Early Impacts of Pay-for-Performance After One Year (NCEE 2014-4019)*, Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2014.

- May, Henry, Irma Perez-Johnson, Joshua Haimson, Samina Sattar, and Phil Gleason, *Using State Tests in Education Experiments: A Discussion of the Issues (NCEE 2009-013)*, Washington, DC: National Center for Education Evaluation and Regional Assistance, 2009. As of August 20, 2015: <http://ies.ed.gov/ncee/pdf/2009013.pdf>
- McDonnell, Lorraine M., and Anthony Pascal, *Teacher Unions and Educational Reform*, Santa Monica, C.A.: RAND, Center for Policy Research in Education, 1988.
- McLaughlin, Milbrey W., "The Rand Change Agent Study Revisited: Macro Perspectives and Micro Realities," *Educational Researcher*, Vol. 19, No. 9, December 1, 1990, 1990, pp. 11-16. <http://edr.sagepub.com/content/19/9/11.abstract>
- Mead, Sara, Andrew Rotherham, and Rachael Brown, *The Hangover: Thinking about the Unintended Consequences of the Nation's Teacher Evaluation Binge.*, American Enterprise Institute for Public Policy Research Teacher Quality 2.0, 2012.
- Miles, Matthew B, and A Michael Huberman, *Qualitative data analysis: An expanded sourcebook*: Sage, 1994.
- Minnici, Angela, "The Mind Shift in Teacher Evaluation: Where We Stand--and Where We Need to Go," *American Educator*, Vol. 38, No. 1, 2014, pp. 22-26.
- Moe, Terry M., "Collective bargaining and the performance of public schools," *American Journal of Political Science*, Vol. 53, No. 1, 2009, pp. 156-174.
- , *Special Interest: Teachers Unions and America's Public Schools*, Washington, D.C.: Brookings Institution Press, 2011.
- Mongeau, Lillian, Emmanuel Felton, and Sarah Butrymowicz, "Stakes for “high-stakes” tests are actually pretty low," *The Hechinger Report*, April 29, 2015.
<http://hechingerreport.org/stakes-for-high-stakes-tests-are-actually-pretty-low/>
- Munguia, Hayley, "New York's Complicated Teacher Evaluation Proposal Is The Exception, Not The Rule," *Five Thirty Eight*, April 14, 2015. As of July 30, 2015.
<http://fivethirtyeight.com/features/new-yorks-complicated-teacher-evaluation-proposal-is-the-exception-not-the-rule/>
- National Board for Professional Teaching Standards (NBPTS), *National Board Standards*, 2014.
<http://www.nbpts.org/national-board-standards>
- National Center on Teacher Quality (NCTQ), *State of the States 2012: Teacher Evaluation Policies*, 2012.
http://www.nctq.org/dmsView/State_of_the_States_2012_Teacher_Effectiveness_Policies_NCTQ_Report
- , *2013 State Teacher Policy Yearbook: National Summary*, National Council on Teacher Quality, 2014a.
http://www.nctq.org/dmsView/2013_State_Teacher_Policy_Yearbook_National_Summary_NCTQ_Report
- , "Teacher Evaluation Policy in California: Where is California in implementing teacher effectiveness policies?," 2014b.
http://www.nctq.org/dmsView/Evaluation_Timeline_Brief_California

- , *NCTQ State Policy: Teacher Evaluation*, 2015. As of July 29, 2015:
<http://www.nctq.org/statePolicy/2014/policyIssueFindings.do?policyIssueId=6&masterGoalId=11&stateId=&yearId=7>
- National Education Association (NEA), *Teacher Evaluation: A Resource Guide for National Education Association Leaders and Staff*, 2011.
<http://www.nea.org/assets/docs/teacherevalguide2011.pdf>
- Nelson, F Howard, and Michael Rosen, "Are Teachers' Unions Hurting American Education? A State-by-State Analysis of the Impact of Collective Bargaining among Teachers on Student Performance," 1996.
- Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E., "Value-Added Modeling of Teacher Effectiveness: An exploration of stability across models and contexts," *Educational Policy Analysis Archives*, Vol. 18, No. 23, 2010. As of August 20, 2015:
<http://epaa.asu.edu/ojs/article/view/810>
- Newton, Xiaoxia A, Linda Darling-Hammond, Edward Haertel, and Ewart Thomas, *Value-Added Modeling of Teacher Effectiveness: An Exploration of Stability across Models and Contexts*, Education Policy Analysis Archives, 23, 2010.
- Pogodzinski, Ben, Regina Umpstead, and Jenifer Witt, "Teacher evaluation reform implementation and labor relations," *Journal of Education Policy*, Vol. 30, No. 4, 2015/07/04, 2015, pp. 540-561. As of 2015/08/03:
<http://dx.doi.org/10.1080/02680939.2014.999827>
- Poole, W.L., "Teachers Union Involvement in Educational Policy Making: Issues Raised by an In-Depth Case," *Educational Policy*, Vol. 13, No. 5, 1999, pp. 698-725.
- , "The Teacher Unions' Role in 1990s Educational Reform: An Organizational Evolution Perspective," *Educational administration quarterly*, Vol. 37, No. 2, 2001, pp. 173-196.
- Rubinstein, Saul A, and John E McCarthy, *Reforming Public School Systems through Sustained Union-Management Collaboration*, Washington, DC: Center for American Progress, 2011.
- Rubinstein, Saul A., and John E. McCarthy, *Teachers Unions and Management Partnerships How Working Together Improves Student Achievement*, Washington, DC: Center for American Progress, March, 2014. <https://www.americanprogress.org/wp-content/uploads/2014/03/Rubinstein-EduReform-report.pdf>
- Ryan, Gery W, and H Russell Bernard, "Techniques to Identify Themes," *Field methods*, Vol. 15, No. 1, 2003, pp. 85-109.
- Sanders, W.L., and J.C. Rivers, *Cumulative and residual effects of teachers on future student academic achievement*, Knoxville: University of Tennessee Value-Added Research and Assessment Center, 1996.
- Sanders, William L., S. Paul Wright, and Sandra P. Horn, "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation," *Journal of personnel evaluation in education*, Vol. 11, No. 1, 1997/04/01, 1997, pp. 57-67.
<http://dx.doi.org/10.1023/A%3A1007999204543>

- Sanes, Milla, and John Schmitt, *Regulation of Public Sector Collective Bargaining in the States*: Center for Economic and Policy Research, 2014.
- Sartain, Lauren, Sara Ray Stoelinga, and Eric R Brown, *Rethinking Teacher Evaluation in Chicago: Lessons Learned from Classroom Observations, Principal-Teacher Conferences, and District Implementation*, Chicago: Consortium on Chicago School Research: University of Chicago Urban Education Institute, 2011. As of August 20, 2015: <http://ccsr.uchicago.edu/publications/Teacher%20Eval%20Report%20FINAL.pdf>
- Sawchuk, Stephen, and Amy Wickner, "States Tighten Disclosure of Teacher Evaluations," *Education Week*, May 23, 2013. As of July 31, 2015: <http://www.edweek.org/ew/articles/2013/05/22/32evaluations.h32.html>
- Steele, Jennifer L., Laura S. Hamilton, and Brian M. Stecher, *Incorporating Student Performance Measures into Teacher Evaluation Systems*, The RAND Corporation, TR-917-CAP, 2010. http://www.rand.org/pubs/technical_reports/TR917.html
- Strunk, K.O., and A. McEachin, "Accountability Under Constraint The Relationship Between Collective Bargaining Agreements and California Schools' and Districts' Performance Under No Child Left Behind," *American Educational Research Journal*, Vol. 48, No. 4, 2011, pp. 871-903.
- Strunk, Katharine O, and Jason A Grissom, "Do strong unions shape district policies? Collective bargaining, teacher contract restrictiveness, and the political power of teachers' unions," *Educational Evaluation and Policy Analysis*, Vol. 32, No. 3, 2010, pp. 389-406.
- Strunk, Katharine O., "Are Teachers' Unions Really to Blame? Collective Bargaining Agreements and Their Relationships with District Resource Allocation and Student Performance in California," *Education Finance and Policy*, Vol. 6, No. 3, 2011/07/01, 2011, pp. 354-398. As of 2013/07/26: http://dx.doi.org/10.1162/EDFP_a_00039
- Superfine, Benjamin M, and Jessica J Gottlieb, "Teacher Evaluation and Collective Bargaining: The New Frontier of Civil Rights," *Mich. St. L. Rev.*, 2014, p. 737. <http://digitalcommons.law.msu.edu/lr/vol2014/iss3/10>
- The Times Editorial Board, "Getting rid of bad teachers," *Los Angeles Times*, October 13, 2013. <http://www.latimes.com/opinion/editorials/la-ed-teacher-firing-20131015-story.html>
- Tyler, John H., *Designing High Quality Evaluation Systems for High School Teachers: Challenges and Potential Solutions*, Washington, D.C.: Center for American Progress, 2011.
- U.S. Census Bureau, "State and County QuickFacts," 2014. As of December 9: <http://www.census.gov/quickfacts>
- U.S. Department of Education, "Shared Responsibility: A U.S. Department of Education White Paper on Labor-Management Collaboration," Washington, D.C., May, 2012, 2012.
- , "ESEA Flexibility," 2015a. As of July 28: <http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html>
- , *Guidance on fiscal year 2010 school improvement grants under section 1003(g) of the Elementary and Secondary Education Act of 1965*, Washington, DC: U.S. Government Printing Office, 2012. <http://www2.ed.gov/programs/sif/sigguidance03012012.doc>

- , "Race to the Top Fund," 2013. As of August 9th:
<http://www2.ed.gov/programs/racetothetop/awards.html>
- , "Teacher Incentive Fund," 2015b. As of June 30:
<http://www2.ed.gov/programs/teacherincentive>
- U.S Department of Education, National Center for Education Statistics, "EDFacts/Consolidated State Performance Report, SY 2012–13," 2015c.
<http://www2.ed.gov/admins/lead/account/consolidated/index.html>
- , "Integrated Postsecondary Education Data System (IPEDS), Fall 2001 and Spring 2007 through Spring 2014, Graduation Rates component. ," 2014. As of August 19, 2015:
https://nces.ed.gov/programs/digest/d14/tables/dt14_326.10.asp?current=yes
- Vachon, Todd E., and Josef Ma, "Bargaining for Success: Examining the Relationship Between Teacher Unions and Student Achievement," *Sociological Forum*, Vol. 30, No. 2, 2015, pp. 391-414. <http://dx.doi.org/10.1111/socf.12168>
- Walton, Richard E., and Robert B. McKersie, *A Behavioral Theory of Labor Negotiations: An Analysis of a Social Interaction System*, 2nd ed., Ithaca, NY: ILR Press, 1991.
- Watanabe, Teresa, "LAUSD ordered to use student achievement in judging teachers," *Los Angeles Times*, 2012. <http://www.latimes.com/news/local/la-me-teacher-eval-20120613,0,3512440.story>
- Weisberg, D., S. Sexton, J. Mulhern, and D. Keeling, "The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness," *Education Digest: Essential Readings Condensed for Quick Review*, Vol. 75, No. 2, 2009, p. 5.
- Wells, John, "Teacher Responses to Pay-for-Performance Policies: Survey Results from Four High-Poverty, Urban School Districts," paper presented at Annual Meeting of the American Educational Research Association, New Orleans, LA, Westat, April, 2011.
- White, Bradford R, Jennifer Cowhy, W David Stevens, and Susan E Sporte, *Designing and Implementing the Next Generation of Teacher Evaluation Systems: Lessons Learned from Case Studies in Five Illinois Districts. Research Brief*. ERIC, 2012.
- Whitehurst, Grover J., Matthew M. Chingos, and Katharine Lindquist, "Evaluating teachers with classroom observations: Lessons learned in four districts.," Brown Center on Education Policy at Brookings Policy Report, May, 2014.
<http://www.brookings.edu/~media/research/files/reports/2014/05/13-teacher-evaluation/evaluating-teachers-with-classroom-observations.pdf>
- Winkler, A.M., J. Scull, and D. Zeehandelaar, *How Strong Are U.S. Teacher Unions?*, Thomas B. Fordham Institute, 2012.
- Zubrzycki, Jackie, "Did D.C. Reforms Mean More 'Irreplaceable' Teachers Stuck Around?," *Education Week*, November 16, 2012.
http://blogs.edweek.org/edweek/District_Dossier/2012/11/did_dc_reforms_mean_more_irrep.html

Appendix A. Characteristics of California School Districts

Table A.1 Demographic Characteristics of California School Districts, 2011-2012

| | Mean | Median | Min | Max | Standard Deviation |
|-------------------------------------|----------|----------|----------|-----------|--------------------|
| Enrollment | 6,393 | 1,906 | 5 | 662,140 | 23,558 |
| % English Learners | 19 | 13 | 0 | 84 | 17 |
| % Free or Reduced Meals | 53 | 54 | 0 | 100 | 25 |
| % Minority | 56 | 54 | 0 | 100 | 27 |
| API Base Scores | 806 | 801 | 564 | 973 | 67 |
| % Schools Making API Targets | 58 | 57 | 0 | 100 | 35 |
| Number of Teachers | 298 | 98 | 1 | 32713 | 1157 |
| Average Teacher Salary | \$64,034 | \$64,203 | \$34,000 | \$101,878 | \$9,251 |
| Average Class Size | 22 | 24 | 0 | 38 | 7 |

Note: data are only available for 968 school districts

Source: California Department of Education, "Student & School Data Files," 2013. As of July 22, 2013: <http://www.cde.ca.gov/ds/sd/sd/index.asp>

Table A.2 Distribution of School Districts in California by Largest Ethnic Group

| Ethnic Group | Number of Districts | Percent of Districts |
|----------------------------------|----------------------------|-----------------------------|
| American Indian or Alaska Native | 8 | 0.8% |
| Asian | 21 | 2.2% |
| Black or African American | 2 | 0.2% |
| Hispanic or Latino | 429 | 44.5% |
| White | 498 | 51.7% |
| (blank) | 5 | 0.5% |
| Grand Total | 963 | 100% |

Note: data are only available for 968 school districts

Source: California Department of Education, "Student & School Data Files," 2013. As of July 22, 2013: <http://www.cde.ca.gov/ds/sd/sd/index.asp>

Appendix B. Interview Recruitment Letter

Dear [Name],

My name is Beth Katz and I am a doctoral student at the Pardee RAND Graduate School in Santa Monica, working toward my PhD in Policy Analysis. My dissertation is focused on how school districts are approaching teacher evaluation reforms. Specifically, I would like to understand how school districts and local teachers unions are working together to develop and adopt new teacher evaluation policies. I intend to summarize the work each district is doing and characterize the interactions and relationships between school boards, superintendents, and teachers unions as those reforms evolve. **I am conducting case studies of three school districts in California, including [name of district].** A summary of my dissertation is attached to this email.

As part of my research, I am interviewing school board members, superintendents, relevant district office staff, and teachers' union leaders and representatives. **I would like to interview you to learn about your district's approach to reforming teacher evaluations,** your role in the development and adoption of teacher evaluation policies, and the relationships between the teachers union, superintendent, and school board throughout that ongoing process. Though there is no requirement for you to participate, **I do hope you will agree to be involved in my study and to participate in a 1-hour interview at a time that is convenient for you.** I believe your insight will be a valuable contribution to my understanding of [name of district]'s efforts. The lessons I learn from the work [name of district] is doing will help policymakers and district leaders who are also interested or engaged in teacher evaluation reforms. I would also be happy to share the results of my three case studies with you.

The information you provide will be kept strictly confidential. I will not disclose your name or the name of your district, nor will I attribute quotes to you or share your responses with anyone. I will only describe general characteristics of your district, taking caution not to include specific information that would reveal your district's identity. The data I collect will be stored on a secured computer and I will be the only individual with access. My procedures have been reviewed and approved by the RAND Human Subjects Protection Committee. Should you have any questions or concerns regarding confidentiality, you may contact me or the Human Subjects Protection Committee directly at 310-393-0411 ext. 6369.

Please contact me by email or phone to let me know if I you agree to participate. If you have any questions or would like more information, please do not hesitate to ask.

Thank you in advance for your participation. I look forward to speaking with you soon!

Appendix C. Stakeholder Interview Protocol

Informed Consent

My name is Beth Katz and I am a Ph.D. student at the Pardee RAND Graduate School, which is based out of the RAND Corporation in Santa Monica. Thank you for agreeing to participate in this interview. The research I am conducting is for my dissertation on how school districts and teachers unions work together on teacher evaluation reforms. I am conducting case studies of three school districts and interviewing multiple individuals within each district. I intend to summarize the work each district is doing and characterize the interactions and relationships between school boards, superintendents, and teachers unions as those reforms evolve.

In my report, I will refer to each district with a pseudonym and will not identify any district by name. In addition, I will not report specific demographic characteristics or other details that would reveal a district's identity. The information you provide will be kept strictly confidential. I will not disclose your name, nor will I share your responses with anyone. Individual responses will be combined with others and reported only in the aggregate. If quotations are used in my report, they will be included only for illustrative purposes and will not be attributed to any individual. At the end of the study, I will destroy any information that identifies you.

I believe your insight will be a valuable contribution to my understanding of your district's reform efforts, and I will be happy to share my results with you at the end of the study. However, your participation in this interview is completely voluntary. You may choose not to participate, decline to answer any question, or stop the interview at any time.

The interview will take about 60 minutes. I will ask you to review the work your district has done to reform teacher evaluations. I will then follow-up with some specific questions about your role and the roles of other stakeholders. At the end of the interview there will be time for you to add anything else you think I should know about. Do you have any questions about my study or the interview?

Do you agree to participate?

With your permission, I will audio record the interview for note-taking purposes only. Of course, let me know if you would like to say anything off the record and I will stop recording and taking notes. I will not include that information in my report.

Do I have your permission to record the interview?

Interview Protocol

[Probes need to be modified so that they are specific to the stakeholder being interviewed.]

I. Background Questions

1. How long have you been in your current position? In the district? In education?
2. In a few sentences, could you briefly describe your current role and main responsibilities?

II. Labor-Management Relationship

I am interested in your district's efforts to change teacher evaluation policies, and the labor-management interactions and relationships throughout that process.

3. How would you describe the labor-management relationship in your district, in general, and also during the efforts to change teacher evaluation policies? Why do you think that relationship exists?

III. Reform Process Overview

4. In your own words, could you please walk me through the efforts to change teacher evaluation policies in your district and how that work has evolved?

Probes:

- a. *Motivation:* Why are there efforts to change the teacher evaluation process in your district? When did those efforts begin?
- b. *Goals:* What were the goals of the district's efforts to change teacher evaluation policies? Were there any formal goal or mission statements?
- c. *Current state of reforms:* How far along is the district in the reform process?
 - i. Has an agreement between the union and district been reached?
 - ii. [If yes]: Has the new policy been implemented?
 - iii. [If no]: Are different policy options being considered?
- d. *Reform features:* Can you tell me about the features of the evaluation policies that were/are being considered? *[Observations, student growth measures, portfolios, student surveys, ratings, etc.]* Why were these features included or not included?

IV. Roles, Interactions, and Relationships

I want to go over how this work happened in more detail.

[May skip questions that were sufficiently addressed in the previous section.]

5. Can you elaborate on...
 - [Superintendent]: Your role throughout the reform process?
 - [District staff]: The district's role throughout the reform process?
 - [Union]: The union's role throughout the reform process?
 - [School Board]: The school board's role throughout the reform process?
6. What roles did the other stakeholders [*school board, superintendent/district staff, teachers union*] play? How were they involved?

Probes:

- a. How did this work to develop a new teacher evaluation system happen? Was there a district plan? Did the district work to develop a new system alone? Was there a multi-stakeholder task force?
 - b. [If there was a task force or committee]: How was the task force formed? Who was on it? What work did they do?
 - c. [For the Superintendent]: How did you decide who should be involved?
 - d. Who served as the leader of the reform efforts?
 - e. When did the union become engaged? Why did they become engaged at that point?
7. How would you describe the relationships and interactions between...
 - [Superintendent]: Yourself and the union/union leader? Yourself and the school board?
 - [District staff]: The district and the union? The district and the school board?
 - [Union]: Yourself and the superintendent? The union and the district? The union and the school board?
 - [School Board]: The school board and the district? The school board and the union?
 ... and what it was like during the process of changing teacher evaluation policies?

Probes:

- a. Over what issues was there agreement/disagreement? Were disagreements resolved?
- b. Were decisions reached in a collaborative way?
- c. [For the superintendent and school board]: Did the [*school board and/or teachers union*] support the district's efforts to change teacher evaluation policies? Why or why not?
- d. [For the union]: Did you support the [*superintendent/district and/or school board*] in these efforts to change teacher evaluation policies? Why or why not?
- e. Did the [*district, union, school board*] have support from other groups [*other relevant stakeholder groups identified prior to the interview, possibly including the administrators union, parents, local politicians, community advocates*]?

- i. Did any other groups oppose these efforts? Why?
8. Why do you think the *[school board, superintendent/district staff, teachers union]* did or did not cooperate or collaborate at various stages in the reform process? What factors made you ultimately implement/not implement the reforms?

Probe for factors that were not mentioned:

- a. District-union relationship
- b. Policies
- c. District history (past reform efforts, status quo)
- d. Culture of the district office, school board, or teachers union
- e. District demographic characteristics
- f. External stakeholders
- g. Significant internal or external events

V. Concluding Questions

9. Do you have any final thoughts or comments that you would like to share?
10. I have looked at *[specific documents]*. Is there anything else you think I should look at that would be informative?
11. Who else would you recommend I speak to?
- a. Could you provide me with their contact information?

Appendix D. Laurel Unified School District Case Summary

Stage 1: Consider Change

Motivation for Reform

Current efforts to change Laurel’s teacher evaluation policies began in 2011. The superintendent, teachers union leadership, and school board members all expressed an interest in improving the district’s teacher evaluation system, as did the new superintendent, who had sat on a teacher evaluation committee in her previous district. The superintendent described it as a challenge to create a robust system that will move teachers’ practice forward, recognize people for excellence, and also address those who need to improve. She found that the existing system did not provide the concrete feedback that teachers needed in order to grow. The teachers union leadership shared her concerns about the existing evaluation system. When the teachers union leadership heard that both the superintendent and some school board members were talking about change, they decided to play an active role, instead of waiting to see what changes might be proposed.

The motivation for initiating reform efforts in Laurel seemed to be both a response to conversations around teacher evaluations and performance-based compensation in the education community and an internal desire to improve a mediocre system. As previously discussed, the *Los Angeles Times*’ publication of value-added scores for elementary school teachers in the LAUSD in 2010 caught the attention of teachers and union leadership. Conversations around teacher evaluations and performance-based compensation were becoming increasingly common in the education community. There was some expectation among district stakeholders that the California legislature would eventually follow other states and adopt new teacher evaluation policy. The union and district leadership wanted to act in anticipation of future state legislation and have a hand in shaping the district’s policies. As one stakeholder said:

And so we were thinking about the fact that we do work well together and isn't there a way to be proactive about teacher evaluation rather than reactive in case something were to come from State or some type of a mandate of doing something differently?

Teachers and administrators in Laurel shared many concerns about the existing evaluation system and wanted to find out if it was effective and actually benefitting teachers. They were concerned about the lack of uniformity in the implementation of the system. Earlier reform efforts were also motivated by a desire to make evaluations more uniform, so that seems to be a persistent challenge. Stakeholders agree that there is still norming to do among principals and other evaluators. Teachers want to be recognized for doing excellent work; however, there was no rating beyond “satisfactory” in the existing system. Though the California Education Code

does mandate the consideration of student performance in teacher evaluations, the district had no strategy for assessing teacher performance based on student achievement. Some stakeholders mentioned that the district was not eligible for certain sources of funding because it did not utilize student achievement data.

Laurel's existing standards-based evaluation system, described in Table 7.1, was the result of a larger district-wide reform effort to increase accountability that began in 2004. Under a previous superintendent, the district contracted with two consultants, who held a two-day workshop on creating a standards-based teacher evaluation system. A joint evaluation committee met to develop a new evaluation tool. Laurel's committee decided to use the six CSTPs as the foundation of the new system. They developed a new instrument, which was piloted in the 2004-2005 school year. To make the process more manageable, each observation focuses on a different group of three CSTPs, rather than all six at once. While the order of standards is set for probationary and temporary teachers, permanent teachers have some discretion; they are allowed to choose two of the three.

The evaluation system was jointly presented to the Laurel School Board by a teacher and district administrator in the fall of 2005.²⁵ In the 2005-2006 school year, the new evaluation tool was used district-wide for probationary and temporary teachers. The evaluation committee continued to meet and after reviewing the pilot and surveying teachers, the new standards-based system was adopted in the contract. The tool itself was last updated in 2008 and as of the 2010-2011 school year, it was used to evaluate all teachers in the district. That original committee's work took a long time and in the end, stakeholders said the evaluation process was improved. Some Laurel stakeholders were disappointed that more changes were not made. For example, the ratings assigned to teachers (satisfactory, needs improvement, and unsatisfactory) were unchanged. They found the work to be worthwhile, but incomplete, which set the stage for the next round of reform efforts.

Decision to Consider Change

The district and union jointly decided the district should form an evaluation committee. This was formally announced at a school board meeting in the fall of 2011. That fall, district and union leaders were also trained in IBB, and contract negotiations began in early 2012. The idea of revising the teacher evaluation policies was discussed in negotiations and that first round of IBB negotiations, which took nearly a year and a half. The union leaders communicated that they wanted an evaluation system that would inform teacher practice, recognize excellence, and improve instruction. Their priorities aligned with the goals expressed by the superintendent.

In early 2012, the superintendent convened an ad hoc meeting of 12-15 district stakeholders, including representatives from the district, union, and school board. One school board member

²⁵ While institutional memory of these reform efforts is limited—few stakeholders interviewed remember much about that work—that committee's progress was documented in school board meeting minutes.

arranged for an outside expert in education policy to attend the meeting and contribute to the discussion. The district wanted to determine if there was enough momentum to undertake the challenging work of designing a new teacher evaluation tool. Though they had recently moved to a standards-based system by adopting the CSTPs, they wondered if there was anything else out there and realized they needed to do more research. The initial meeting was rather informal and they did not outline any clear next steps. However, the district did name “reviewing the teacher evaluation process” as one of its goals for the 2012-2013 school year, along with the broader goal of “collaboration” and continuing with IBB. The original ad hoc group only met once, but a few leaders, including the superintendent and union president, met a few more times that year.

As noted earlier, union leaders were proactive in these efforts; they did not want to wait for a new system to be imposed upon them and to be put in a reactionary position. They met over the summer to discuss the role they wanted to play in changing district policy. No work had been done since the ad hoc committee met earlier in the year and they realized they needed something more organized if they were going to make progress. At the beginning of the 2012-2013 school year, the union president and one representative approached the superintendent about finally getting the evaluation committee off the ground.

Committee Recruitment

The district and teachers union chose their own representatives to sit on a new evaluation committee. The union sent an open invitation to all teachers in the district to join the committee and one individual recalls being personally approached by a union representative about participating. Participation in the committee was voluntary and committee members were compensated by the district for their work. Only five teachers volunteered to participate and the union leadership accepted all of them. These individuals were motivated by a personal interest in teacher evaluation reforms and felt strongly that this work was important. As one committee member said of the group, “We all had like some personal drive to do something, like to make a difference, to do something. We all, I think were motivated to be there.” The superintendent and Assistant Superintendent of HR recruited the district representatives. They wanted to include individuals from different offices, as well as a principal and a BTSA provider, since they had experience in supporting teachers. They also thought it would be valuable to have someone who had worked on teacher evaluations in another school district. The Assistant Superintendent of HR took the lead for the district. Though he would sit on the committee, he maintained that it would be important for the teachers to own the process and did not want to be the leader of the committee. A veteran teacher, who was a union representative, volunteered to take the lead. The group that was formed was smaller and more organized than the initial ad-hoc group that met earlier in the year and would meet for the first time in October of 2012.

Stage 2: Set Goals

The evaluation committee first met in October, 2012, and continued to meet monthly after school, for a total of 10 times in the 2012-2013 school year. Laurel had approximately 12 participants at some point, nine individuals considered to be the core committee members. A union representative gave a brief presentation to the school board in November 2012 to introduce the committee to the district. The presentation also discussed the broader goals of building a community based on collaboration and professionalism. No school board members were directly involved in the teacher evaluation work. At least one school board member, when interviewed in 2014, was still unaware that an evaluation committee had been convened at all.

The full committee consisted of approximately 12 individuals, but meetings were consistently attended by four to five teachers, at least one representative from the superintendent's office, and at least one school administrator. There was a good mix of both older and younger teachers. One stakeholder described the union members involved as "dedicated professionals who want to better their craft, and see evaluation as one tool to do that."

From the beginning, one union representative assumed a leadership role. She facilitated the meetings, sometimes along with the Assistant Superintendent of HR. While the superintendent joined only a few meetings, acting mostly as an advisor, her presence reaffirmed the district's support for the committee. That did not go unnoticed among the other participants. Having the superintendent directly participate lent credibility to the effort, increased buy-in of the key players, and signaled it was a district priority and that their work would actually have an impact. Some stakeholders noted that district task forces, of which Laurel has had many, often end up seeming pointless.

The mission of the committee, while not precisely defined, was essentially to evaluate the existing system and set goals for improving it. This was articulated in the committee's final letter of recommendations, which was drafted at the conclusion of their work:

We wanted to be pro-active and make decisions based on meaningful research.
Our goal is to create a meaningful evaluation instrument that provides feedback
to members to improve instructional practices.

The superintendent had her own idea of what the eventual outcomes of this work would be. She hoped that the committee would judge the current process, identify what aspects were good and bad, examine the idea of a career ladder with more opportunities for teachers, and improve calibration to make sure the system is consistently applied, resulting in a meaningful process for both the teachers and administrators.

In Laurel, the committee began its first meeting by talking about what motivated each of them to volunteer to serve on the committee in the first place. The committee saw itself as a research group that would provide specific recommendations, but would not actually design a new evaluation system. The committee did not have the authority to make policy changes, which would need to be negotiated. Their work was concurrent with, yet independent from ongoing

contract negotiations. The Assistant Superintendent of HR sat on both the evaluation committee and negotiations team, and was able to communicate between the two. At the beginning of the year, the negotiations team prepared a list of 16 shared interests related to teacher evaluation, which they asked the committee to discuss over the course of the year.²⁶ For example, one “shared interest” was that they wanted to differentiate excellent teachers from satisfactory teachers. There was communication between the negotiations team and committee members throughout and the committee’s final product was a letter with recommendations sent back to the negotiations team.

The committee’s work was described as “exploratory,” with the goal of “learning first” before making changes to the existing evaluation system. One stakeholder described it as a “mini think tank” and an “intellectual exercise.” There was no hard deadline for designing a new system, nor a palpable sense of urgency to implement change. While the group lacked focus at the initial meeting, they became more organized at future meetings. They set meeting agendas together and assigned tasks to complete between meetings. The group would pick two to three peer districts to research and divide them up among members on a voluntary basis. Individuals would collect relevant readings and distribute them to the group prior to meetings. Members commented that individuals did indeed complete their tasks and come prepared to meetings, for the most part. The meeting format included a presentation by one member, followed by an open discussion. They created a spreadsheet that described policies in 10 districts and CMOs across the following characteristics:

- Relation to merit pay and tenure;
- Career pathways;
- Professional standards;
- Peer involvement/review and support;
- Assessment instruments and contribution to overall rating; and
- Frequency of classroom observations.

Documents, including the readings and notes from the meetings, were uploaded to a document sharing site that was accessible to all teachers and other district stakeholders, not only to committee members.

The group began by diagnosing the existing system – identifying its strengths and weaknesses. They generated a “wish list” for what evaluation system features would be valuable to them. They identified several weaknesses in the existing system that needed to be addressed. Both teachers and district administrators agreed that uniformity in evaluations was essential, and that the current system was deficient in that area. While principals use the same documents, they may use them in different ways and had never been normed. There was also agreement that the existing three-level rating system and its implementation was not ideal. Teachers wanted to be recognized for excellence, for being “beyond satisfactory.” Interviewees also noted a lack of

²⁶ This process is part of IBB negotiations.

support for struggling teachers. While there was a procedure teachers follow if they did not reach “satisfactory,” stakeholders described it as “hoops” to jump through. They believed teachers could improve even more if there were better tools and support available to them. The group looked at models of PAR in other districts that were used to help teachers grow. The group also agreed that it would be important to have concise and clear language. They found some of the evaluations systems from other districts to be too wordy.

Laurel’s committee’s work was mainly focused on the research task – reviewing literature on teacher evaluations and finding examples from other districts – as they considered themselves to be a research group. Individuals were assigned districts to research and then would present their findings to the group at meetings. Committee members created a template to compare the policies of ten districts and CMOs across several characteristics, which motivated discussion. The characteristics they reviewed included:

- Relation to merit pay and tenure;
- Career pathways;
- Professional standards;
- Peer involvement/review and support;
- Assessment instruments and contribution to overall rating; and
- Frequency of classroom observations.

One committee member described the discussion that followed:

When you're looking at it, this stuff's interesting. Looking at charter schools, and looking at Green Dot, these KIPP schools, they are very interesting in a way that their whole philosophy is. And so it just made for a rich discussion because, what they do is so, has some really unique characteristics. You might value them, you might not, but it was definitely interesting to have looked at them, to choose such a variety of different types of systems.

The group used this format to identify features in other the other systems that they might want to incorporate in their own. Laurel’s committee, as well as individuals in other districts, also noted reading articles by Charlotte Danielson and Linda Darling-Hammond.

The group found many commonalities up front, but also made a tacit agreement to avoid certain controversial policy options. While the group discussed the potential role for student achievement data in an evaluation system, the idea of using value-added measures to compute a teacher’s rating, as well as the possibility for performance-based compensation to be associated with evaluations, were off the table. Committee members also noted that the discussion of student achievement had to be tabled given the uncertainties inherent with the transition to the Common Core standards and changes to California’s standardized tests. The district and teachers were both uncomfortable with the idea of using student test scores to evaluate teachers during that period of change. While specific measures of student achievement may ultimately be an

important reform feature to discuss, by tabling that discussion, it allowed the committee to focus on other topics and continue with their work.

All participants interviewed agreed that the committee's work was collaborative and that they were able to move forward as a result. The teachers on the committee were motivated individuals with a pre-existing interest in this work. Because the district is relatively small, several committee members already knew each other. Some district administrators had previously worked with some of the teachers when they were school administrators. One stakeholder said, "If anyone can do it, we can." In addition to the district being relatively small, the group was small enough to sit around one table, which aided communication. Being able to sit around one table and see all committee members contributed to positive interactions. During the meetings, people were respectful and good at sharing. However, they were often guarded during discussions and were careful not to challenge each other. One stakeholder noted that while the teachers seemed to have strong opinions, especially around certain controversial issues, the district and school administrators tried to maintain a neutral tone during meetings. The absence of conflict during meetings may also be attributed to the fact that the group did not begin the work of defining and debating discrete policy options.

Laurel posted all their notes and resources on a district document sharing website, which could be accessed by any employee. There were informal conversations between teachers and union representatives and throughout the process, union leaders posted updates on the union website. The lead union representative also presented at union meetings. At the end of this stage, the entire committee distributed a final letter of recommendations to the district community and presented at a school board meeting.

Current State of Reforms

At the end of the 2012-2013 school year, the committee completed the goal-setting stage and sent a summary letter with their recommendations to the entire district community: the school board, administrators, negotiations team, and all teachers. However, the extent to which the district community was aware of this work is unclear. One union leader did note that some teachers would ask her questions about the committee's work and that they wanted to know if they were discussing merit pay and the use of student achievement data. However, as discussed earlier, not all school board members were aware of the committee's work or its final recommendations. There were no regular updates on the committee's progress at monthly school board meetings and the topic of teacher evaluations may no longer be on their radar.

The committee's final recommendations included the following:

- Expand the purpose of the evaluation system to focus on growth, support, and identifying excellent teachers.
- Adopt the Danielson framework to replace the CSTPs and replace the existing three-level ratings system with four levels.

- Increase the frequency of observations and allow teacher peer leaders and other school administrators to conduct classroom observations and train all teachers and observers.
- Incorporate multiple measures and limit the use of student achievement data to goal-setting and reflection, while not incorporating it into a teacher's overall rating.
- Develop versions of the evaluation system for certificated staff who are not based in the classroom.

Their recommendations were somewhat vague. Though they recommended the use of “multiple measures” of teacher effectiveness, they did not specify which measures should be used. They suggested an increase in the frequency of classroom observations, but did identify a desired number. This lack of concrete details speaks to the unfinished state of their work.

The committee also identified three next steps for continuing the process:

1. Review and modify the Danielson framework for use in a new evaluation tool.
2. Design options for a pilot evaluation system.
3. Have the negotiations team revise contract language.

However, there has been no progress since the committee last met in May, 2013. Their recommendations went back to the bargaining table at the same time negotiations were in-progress; the timing of this work with the current contract negotiations cycle was misaligned. Evaluations were not being discussed at the bargaining table and the next cycle of negotiations would not commence until January, 2016. Further, the committee did not propose a timeline for continuing the process. The district seems to have lost the momentum that initiated these efforts. Because the committee was seen as “exploratory,” there was never the expectation that their work would move beyond goal-setting. One stakeholder said that the current state of reforms can be described as “so far, so good.” In other words, there has not been any real conflict yet because there are no concrete proposals on the table. While the committee discussed many features of teacher evaluations, they concluded without tackling what one stakeholder referred to as “the hard stuff.” The group was not ready to take on some of the more difficult and controversial topics, including the role of student achievement data. The transition to Common Core and uncertainty about new state standardized tests served as a reason to table the conversation about incorporating student data in teacher evaluations. Still, despite the waning momentum, stakeholders were proud of the work they did. A new union president took office in 2014, after the teacher evaluation committee last met. She therefore has not participated in the teacher evaluation reform work. How the most recent change in union leadership will affect the labor-management relationship remains to be seen.

Stakeholder Relationships

The state of the labor-management relationship in Laurel prior to the initiation of the teacher evaluation reform efforts was the most positive of the three case study districts. When their work terminated after the goal-setting stage, that relationship had only improved. The new

superintendent, who joined the district right before they began to consider change, made labor-management collaboration a priority. One of her first initiatives was to transition the district to IBB and she articulated collaboration as a district-wide goal. In that first year, she was able to foster a positive relationship with the union president and they communicated openly. Stakeholders described a general “collaborative vibe” in the district.

The interest in teacher evaluation reform in Laurel came from both sides and the decision to form an exploratory committee was mutual. Many stakeholders interviewed were not sure who proposed the idea for the committee in the first place. Interactions among members of the teacher evaluation committee were collegial and positive, as is characteristic of the collaborative relationship type and also reflective of a broader district culture. Teachers led the committee and did not express concern about the district’s intentions with respect to teacher evaluation policy changes. In these first two stages of the reform process, the stakes were low for teachers in Laurel. The committee was not tasked with designing a new system, so implementation of a new teacher evaluation and its potential consequences were not an immediate concern. While this was an enabler of difficult conversations, it could also have been a barrier to progress. Laurel’s committee never moved beyond the goal-setting stage. In the future, they will have to grapple with concrete policy options and address issues likely to uncover differences in opinion. It is possible that the priority placed on working together and getting along may have made them overly cautious. Walton and McKersie (1991) observed that positive attitudes between labor and management can actually make it difficult to address emerging conflicts. Koppich and Kerchner (2000) also noted that while a trusting relationship has benefits, trust alone will not produce reforms; the parties must be willing to accept change. To move forward with reform, stakeholders will have to be willing to rock the boat and put their collaborative relationship to the test.

Appendix E. Juniper Unified School District Case Summary

Stage 1: Consider Change

Motivation for Reform

Juniper stakeholders recalled hearing a great deal of conversation around the high-profile education policy reforms that were spreading nationwide in the mid-2000s, including the work of Michelle Rhee in D.C. Some say that the Los Angeles Times' value-added scores "scandal" may have contributed to the pre-existing distrust that the teachers union had for the district. While there was no direct pressure from the community to overhaul the district's teacher evaluation system, there did seem to be a growing support for a performance-based compensation system for teachers. One individual recalled reading several letters to the editor in local newspapers, written by parents who supported such policies.

The district was simultaneously engaged in other reform initiatives related to curriculum and technology. As the district began planning to implement the newly adopted Common Core standards in 2010, there was a desire to align PD with those standards. Around that time, the Juniper School Board began a process of setting annual goals and expressed an interest in an improved performance evaluation process that would be tied with professional development and coaching. The existing teacher evaluation system in Juniper, as outlined in the 2009 CBA, was not considered to be a very effective tool. They were concerned about a lack of consistency in implementation. It was not standards-based and did not serve as a professional growth tool. Additionally, it did not distinguish between teachers who met or exceeded expectations, having only three ratings categories: "satisfactory," "needs improvement," or "unsatisfactory." Some stakeholders wanted to see the process recognize teachers for excellence. Lastly, it seemed as if teachers and administrators were just "going through the motions" with the existing evaluation system.

Decision to Consider Change

When Juniper's current superintendent took office in 2010, she created several district committees around a number of initiatives. She had worked to design a new teacher evaluation system in her previous district and knew that teacher evaluation reform was a school board priority in Juniper. The superintendent approached Union President #1 about forming a committee to improve the district's teacher evaluation tool and they agreed to initiate the process. Stakeholders did agree that the existing evaluation tool was not helping teachers improve their practice. As one individual recalled:

And we also agreed that the current evaluations were ineffective and inconclusive, and kind of irrelevant toward anyone and their teaching. It was some motions that everybody went through, and so impractical on every level.

The formation of the committee was officially agreed upon in contract negotiations.

While district and union stakeholders shared many concerns, some teachers remained skeptical of the district's intentions. They were concerned that the district really wanted to change the teacher evaluation system in order to weed out non-tenured teachers. Others wondered, given that they were a high performing district, if they really needed a new evaluation system or if the district was exploring those reforms out of political necessity.

Committee Recruitment

The Juniper Teachers Association invited all teachers in the district to join the first teacher evaluation committee, but there was little response. Union President #1 said she looked for people with diverse ideas about teacher evaluation. Union leaders ended up choosing five representatives: UP #1, a veteran National Board certified teacher, the union negotiations chair, an elementary teacher, and a secondary teacher. The district brought in the superintendent, who served as the leader, as well as a high school principal and an elementary school principal. One school board member and an interested parent offered to serve on the committee. However, it was decided that having them on the committee could prove to be counter-productive, and that it was best to leave the decision-making to teachers and administrators, so that they would not feel like they were being told what to do. The first committee would begin meeting at the end of 2010.

Stage 2: Set Goals

Juniper's first teacher evaluation committee met several times during the 2010-2011 school year. During that period, committee members took a trip to a peer district in another state to learn about their teacher evaluation reforms and met several times, sometimes for full-day discussions. While the superintendent and UP #1 co-facilitated the committee, the superintendent seemed to be driving the work. The committee spent some time upfront to discuss the purpose of the new evaluation tool they hoped to design. They agreed to move toward a standards-based system. The district did not want to advocate for a particular set of standards. However, they were perceived to be supporting the NBPTS. Some teachers wanted to continue to use the ESLRs that individual schools in the district had already created as the basis for their professional standards. After some debate, the committee determined to develop its own six standards.

The union was particularly concerned about the purpose of a new teacher evaluation system. They did not want the tool to be used as a "gotcha" for catching teachers doing something wrong, which could have consequences for their job security. They wanted to make sure that there were protections built-in to guard against unreasonably negative evaluations, which could

possibly result from parent complaints. While the superintendent maintained that the sole purpose of the evaluation, from her perspective, was to promote professional development, the union continued to express their concerns during conversations at committee meetings. After significant discussion, they eventually agreed that their goal, as a committee, would be to create a professional growth tool that would improve teacher practice. They also acknowledged the importance of working with principals to ensure that evaluations were effective and that they would be able to differentiate among teachers. Still, some teachers remained skeptical of the district's intentions.

The district made a conscious decision not to push for the inclusion of student achievement data. While the superintendent said that the use of student data is an area of growth for the district as a whole, they were not in the position to be using student data to evaluate teachers at that time. In May of 2011, the *Los Angeles Times* released value-added data for more teachers in the LAUSD, which fed Juniper's teachers' existing concerns about the use of student achievement data.

Juniper's superintendent brought in resources for the committee to review, including examples of teacher evaluations in other districts and related research, though the group did not devote much time to looking at examples from other districts. Committee members would also do some reading on their own, then come together for half-day or full-day discussions. Juniper spent less time on research activities than the other two districts, as they progressed more quickly through the goal-setting stage. However, this was the only committee which took a trip to visit another school district that had developed a new teacher evaluation system that was fostered by excellent labor-management relationship. Juniper's committee did engage in discussion about what the purpose of their evaluation system should be, which was difficult at times. Though district administrators stressed that the evaluation system would not serve as a "gotcha" for catching teachers doing something bad, some teachers remained skeptical of the district's intentions. Concerns about the district's intention persisted throughout the reform process. The group also discussed the alternative evaluation already written into their contract. Since few stakeholders had opted to participate in the alternative evaluation, many did not understand its purpose or related procedures.

The goal-setting stage was rather brief in Juniper, when compared to the other case study districts. In the superintendent's first year in the district, they went from beginning to consider change to forming a committee, clarifying the purpose of an evaluation system, and agreeing to continue their work. Stakeholders said that people knew the committee existed, because they were all invited to join. As one individual said:

They knew that it was there, but they didn't know what the committee was trying to do. And there were some broad strokes, revising the evaluation process, so that the evaluation process was going to be more stringent and we had also talked publicly about it being a professional development tool. But, I think that's, it's kind of vague, right? In the way that that looks?

There did not seem to be as much direct communication about the committee's work during the goal-setting stage.

Stage 3: Define and Discuss Options for Change

Juniper's teacher evaluation committee continued to meet over the 2011-2012 school year. Committee members would read independently outside of the meetings and then come together for discussions. There was some underlying and persistent distrust among the committee members, although it did not paralyze the group. The most difficult conversations were around developing the standards and determining the cutoff points between the different levels on the evaluation rubric. It was especially challenging to differentiate between the criteria for "meeting" versus "exceeding" standards. The committee also discussed the possibility of using peer review. However, the union was concerned that it could lead to divisiveness and unnecessary competition among teachers. They did not like the idea of establishing a hierarchy and putting mentor teachers in an evaluative capacity.

The committee eventually agreed that there would be six standards. One of the compromises was that teachers would not be evaluated on all six standards in one evaluation cycle. All teachers would be evaluated on the two standards related to teacher professionalism and then two others – one chosen by the teacher and one chosen by the administrator. This was an area that reignited debate in Juniper when the composition of their teacher evaluation committee changed. Teachers were concerned that there were too many standards and that it would be difficult to remember what they were.

In Juniper, the district wanted three ratings levels, but the union preferred to have only two. They compromised by including three levels on the rubric – "does not meet standard," "meets standard," "exceeds standard" – with two sub-ratings in each category. In other words, a teacher would be rated on a six-point scale, but receive one of only three ratings. This was a departure from the status quo, as the existing system had two categories that fell below "meets standard" and did not distinguish between "meets" and "exceeds." However, the final evaluation rating would still be binary, either "meets or exceeds standard" or "does not meet standard." There were no specific instructions for translating the ratings on the sub-standards to the overall rating for a standard, as there was in Sumac. In writing the rubric, participants found it particularly difficult to define the boundary between "meeting" and "exceeding" a standard. There were a few exceptions to the three-level rating system for sub-standards that were difficult to "exceed," such as submitting paperwork on time. This was an area where Juniper's second committee decided to change language that was agreed upon by the first committee.

In Juniper, the use of student achievement data was prescribed by Standard 4. One stakeholder in Juniper noted that the use of student data will continue to be an area of growth for the district because they do not yet have the tools in place to effectively utilize that data.

Teachers in Juniper did not like the idea of peer review; they felt that it would put teachers in an awkward situation. As one committee member described:

The peer to peer evaluation was an issue, like somebody being the mentor creating divisiveness among the teachers or competition among the teachers that they didn't think was important, or necessary, or appropriate. And the idea that these mentor teachers, would they be involved in or have any sort of say in the evaluation of these teachers? But they didn't want to create a hierarchy among the teachers.

They were not supportive of having teachers conduct classroom observations or complete surveys about their peers. Juniper also included an element of teacher choice in its pilot evaluation tool, allowing teachers to choose one standard as their own professional growth goal. Their administrators would choose the second goal.

Stage 4: Come to Agreement

Juniper's first teacher evaluation committee was able to come to agreement on a rubric and other forms for the pilot evaluation, and they were proud of that achievement. However, when a draft was released to the district community, one committee member did withdraw support. They sent the documents they created to the negotiations team, which would have to officially sanction the pilot when a new round of bargaining began in March of 2012. At the same time, there was a change in union leadership. In the fall of 2011, UP #1 decided to not run for reelection. The new candidates ran unopposed and UP #2 was elected in 2012. The union actually struggled to fill all their representative positions. When the new leadership took office in the summer, labor-management relations became more difficult.

In Juniper, the terms of the agreement drafted in the first teacher evaluation committee had to be formally adopted during contract negotiations before the pilot could begin. Negotiations began in the spring of 2012 and continued into the fall, eventually requiring mediation. Despite the contentious environment, a tentative agreement was reached in September and soon after, the negotiations team drafted an MOU regarding the teacher evaluation pilot. The MOU included a timeline for running a pilot of the new teacher evaluation system. The pilot would start in 2013-2014 and continue in 2014-2015. As part of the agreement, the pilot will only be mandatory for probationary and temporary teachers; permanent teachers would participate on a voluntary basis. The documents created by the first committee would be used for the pilot, unless both sides later agreed to make changes. The MOU also established the second teacher evaluation committee, which would oversee the pilot. That committee would consist of four teachers and four district representatives and be co-chaired by a district representative and union representative. The MOU only applied to the teacher evaluation pilot. There were no changes to the actual contract language regarding teacher evaluation and the MOU did not ensure that the pilot system would eventually be adopted district-wide. Future negotiations would be needed to change the contract language.

One stakeholder thought that an agreement was inevitable given the investment that had been made over the previous two years:

I think because of the two years of work that the committee did. There had to be results, right? There had to be something. And so I think that's why the MOU was created, it was because they'd done two years' worth of research and in creation of these documents and I would agree that yes, we'd go ahead and pilot it.

The union leadership found the tentative agreement palatable for several reasons. First, they received a salary increase in exchange for moving forward. Because of that, some teachers were resigned to accept the pilot, as one individual said:

The only reason this particular evaluation system is in place right now as a pilot, is because we had to take it. They tied the pilot evaluation system to the raise we got last year. They would not give us the without the agreement to do the pilot evaluation. This was absolutely an arm-twist to make us do it.

Second, the pilot was not mandatory for tenured teachers. Thus the union was not concerned about any immediate negative consequences that could potentially be tied to the results of a teacher's evaluation. Third, there would be continued mediation, as part of contract negotiations, before a permanent system would be adopted. In that sense, it was a "low stakes" decision for union leadership; they knew they could still negotiate a different policy in the next round of negotiations, if they were not happy with the outcome. As one union representative described:

Well, one of the reasons why we agreed, and I don't know if it is a great process, was because it didn't matter to us, because it was a pilot, and we knew that after a couple of years, we could negotiate what we wanted to anyway.

Setbacks

The change in union leadership posed challenges. When the new union leadership took office in 2012, they decided to change the composition of the teacher evaluation committee, retaining only one representative from the previous committee, and dropped all other joint district-union committees. The union representative who served on the first committee was chosen to co-chair the second committee with the superintendent. The outgoing union president, who was on the first committee, was not invited back by the new union leadership.

Though the first committee had agreed on the procedures and documents for the pilot evaluation, the second committee reopened discussions that had been previously settled, essentially setting the committee back to the "define and discuss options" stage. As one committee member described, "Then we inherited their work, and then we went to work on trying to optimize what they had come up with." They kept some of the forms that were previously developed, but spent a lot of time reworking the language of the rubric. They wanted to go in a different direction in some areas, as one individual commented:

We spent almost all of our time on the rubric, and redefining that. And some things the union at that time just said, "We... It's too vague for us. So we got to take the whole thing out." So... And in some cases, it was a good call. I mean it was too... Maybe redundant or it wasn't clear enough and so... Other cases, I disagree, but it wasn't worth shooting the whole thing down for it. So, it became a slightly scaled down version because of all that.

The union continued to express concerns about calibration among evaluators and consistency in the application of the evaluation tool, as articulated here:

And so, one of the things that we were working towards is making sure that the administrators know the tool. They know how to use the tool. They understand the flaws of the tool, they standardize the use of the tool, and essentially, make sure that the administrator has the capacity to use the tool and not misuse the tool.

The union representatives wanted fewer standards and more constraints; they also wanted to address areas they found vague or redundant. The district wanted to keep the standards, but also maintained that they did not expect teachers to master all of them at once and in one lesson. The committee eventually agreed to reduce the number of sub-standards.

The second committee in Juniper found the professional practice standards concerning for several reasons. These standards included elements related to how parents perceive teachers, as well as teacher participation in professional development workshops and conferences. One individual thought it was unfair to assess teachers on participation in conferences that they do not get paid to attend. According to Juniper's rubric, teachers would receive a "does not meet standard" rating on professional development for the following criteria:

Never or rarely participates in workshops, conferences, activities sponsored by professional organizations, or district.

Another individual noted that not participating in conferences does not indicate that a teacher is uninterested in growing as a professional; they might not feel that the opportunities presented to them are worthwhile or meet their needs. Further, one individual felt that it would be unfair to judge teachers on their participation in district-created PD if teachers are not involved in planning those opportunities.

While evaluators might be able to track participation in workshops, some participants felt that they would not be able to accurately assess the extent to which teachers share with their colleagues; two of the professional development sub-standards in Juniper address collaboration with colleagues. As one individual said:

So "doesn't share materials and experiences with colleagues." "Does not plan, evaluate or reflect with colleagues." How do you know that I'm doing that?

Another individual took issue with the expectation that teachers, including veteran teachers, would have to change their practice every year:

"Seeks out and develops new methodologies to continually enhance teaching with regard to changes in student population and the education pedagogy." So

you have a lot of polysyllabic words in there, so it must be super smart. But why do you need to constantly change what you're doing? Change does not effectiveness make. So they're constantly including these buzz words that are all part of some educationalese [sic] and some designed to create the appearance of effectiveness, but which don't actually have anything to do with effective teaching. Why do I have to change what I do every year in order to be considered effective?

The sub-standard referenced here is labeled “commitment and tenacity,” which she found to be vague.

The professional practice standards in Juniper also addressed communication with families, and some teachers took issue with the expectation that they frequently communicate concerns to families and use technology to do so. One individual commented that in their high-performing district, many parents complain when their children do not receive ‘A’ grades and that dealing with those complaints is time-consuming. This echoes other concerns about parent complaints influencing teachers’ scores. This is likely a greater issue for secondary teachers who have more students, and thus more parents to communicate with, than elementary teachers. Teachers were also concerned about how they would be judged on the use of technology, which also comes up in another standard. One stakeholder described those concerns:

So technology in general was a huge issue. And part of the problem was the idea that teachers are going to be evaluated on their use or lack of use of something that they haven't been trained to use. And that's huge. And with inadequate training, and then you're saying you have to use this online grade book to somebody who's never turned on a computer. And that's valid, I think a really valid point that they made.

While Juniper has engaged in initiatives to implement more technology in classrooms, not all teachers may be prepared to utilize that technology.

The character of labor-management interactions was different in the second committee. Meetings were more formal and were run like negotiations; the first committee had been more collegial and less formal. They did eventually come to agreement on the rubric and related documents at the end of the 2012-2013 school year. However, it was described as a “reluctant agreement,” rather than an accomplishment to celebrate. One individual described the difference between the two committees’ attitudes:

Well, particularly in the first committee, we all felt very good about it. We were pretty celebratory in finishing. In the part two, the phase two of the committee, we collaborated and had an agreement and shook hands at the end. But it was kind of, "Okay, I guess we can all agree on this," as opposed to an enthusiastic... It was a reluctant agreement, as opposed to an enthusiastic agreement.

Though there was less enthusiasm and an increase in resistance, Juniper was still able to move forward with implementation of the teacher evaluation pilot.

Stage 5: Begin Implementation

In July, 2013 – the summer before the pilot evaluation was set to begin – Juniper experienced more leadership turnover. The district hired a new Assistant Superintendent of HR and a new union president (UP #3) took office. The Assistant Superintendent of HR took the reins on the evaluation pilot and also began working on new administrator evaluations. Contract negotiations, which began in the spring, continued into the fall and were fairly contentious. There was significant conflict between teachers and the administration at a high school over several issues, including the implementation of the Common Core standards. Though the district was able to move forward with the teacher evaluation pilot, labor-management relations were strained.

Juniper's new teacher evaluation system was based on the NBSTPs. Teachers in Juniper were rated as in "meets or exceeds standard" or "does not meet standard." However, on the rubric used for classroom observations, administrators in Juniper rated teachers on a scale of one to six for each of the sub-standards. Temporary and probationary teachers were observed twice and permanent teachers were observed once, with observations lasting about thirty minutes. The district required a goal-setting conference with the teacher and administrator in the fall, and a post-observation conference to debrief the observed lesson.

There was some concern about whether or not principals were prepared to conduct the new teacher evaluations. Union leaders met with the superintendent at the beginning of the year and insisted that every principal should teach a lesson and get the experience of being evaluated by the new tool before using it. The district agreed and the superintendent also offered to teach a lesson and have a union leader evaluate him using the observation rubric. The Assistant Superintendent of HR conducted a two-hour training on the pilot for principals and supported them throughout the year. He did a collaborative evaluation with each principal where they would watch a full lesson, compare ratings, and then debrief with the teacher together.

The pilot was rolled out as a joint union-district initiative, though the Assistant Superintendent of HR really took the lead on implementation. He held trainings in September and all teachers – not only those participating in the pilot – got paid to attend. The Assistant Superintendent of HR went to every school to introduce the pilot and explain how teachers could participate. The newly elected UP #3 joined him at most of the schools, and the superintendent came along at a few. They reviewed the details of the pilot and who would be required to participate. The union conducted separate short information sessions for teachers and communicated with their members through email. Juniper offered paid trainings to any teacher willing to attend, whether or not they would participate in the pilot.

Not many permanent teachers were opting-in, and the district felt that the union leadership was discouraging participation. On the other hand, union leaders felt that teachers were receiving misinformation about who was required to participate. At one school, the principal erroneously told all the teachers that they had to be evaluated that year. The district and principal later admitted there was an error in record-keeping and that not all teachers were in the evaluation

cycle that year. Teachers at one school also complained that the principal was doing too many “drop-in” observations and were skeptical of the purpose of those visits. In total, about 25% of Juniper’s teachers participated in the first year of the pilot. About two-thirds of participants were temporary and probationary teachers, all of whom were required to participate, and the remaining third included the permanent teachers who chose to opt-in. At one school, no permanent teachers opted-in. Participation grew only slightly in the second year.

While the teacher evaluation last met in the fall of 2013, the district continued to monitor and evaluate the pilot. The superintendent and Assistant Superintendent of HR had many conversations with participating teachers; they held two-hour “think tank” discussion with the teachers. The district also conducted a survey of all participating teachers and principals. They received mostly positive feedback. Teachers said they would recommend the new system to their peers and also felt like the process was collaborative and that they were receiving targeted feedback aligned with their goals. Principals also gave positive feedback, although they were concerned at the amount of time the evaluations were taking.

In early 2014, the school board requested a board workshop on the teacher evaluation pilot. The Assistant Superintendent of HR led the meeting and went through all the documents and had the entire group watch a lesson through the lens of one of the standards. The superintendent and a principal joined him in presenting a demo on how a post-observation debrief conversation should be conducted. Lastly, the district presented the results of their surveys and conversations with participants. The school board was very supportive of this work. Though the workshop was open to anyone interested, it was not well-attended by teachers. It is unclear how much attention was being paid to the pilot by those who were not participating.

The superintendent and Assistant Superintendent of HR held a meeting in January of 2014 with all participants to solicit feedback, and conducted a survey. The Assistant Superintendent of HR also received feedback from administrators after joining each principal in conducting and evaluating a classroom observation. Average ratings in the first year of the pilot fell between ‘4’ and ‘5,’ on standards 1 to 4, with no teacher receiving the lowest rating of ‘1.’ Elementary teachers received higher scores than secondary teachers. Teacher strengths identified in the pilot included their grading systems. There were also areas of growth, as teachers scored lower in the area of student assessment. Interestingly, the results from the pilot suggested that, if anything, ratings were biased upwards, with more teachers receiving high ratings than low ratings.

Teachers in Juniper provided some positive feedback. According to survey data collected by the district:

- Nearly 8 out of 10 participants found the pilot to be more effective than the pre-existing evaluation process.
- Almost 4 out of 10 felt overwhelmed by the number of elements in the standards they were being evaluated on.
- Less than 5% of teacher participants thought the rubric aimed to penalize teachers.
- 86% would recommend the system to other teachers.

They found that the new evaluation promoted more meaningful conversations between teachers and principals. They also felt more involved in the evaluation process itself because they were able to choose one of the standards on which they would be evaluated. However, there was some less positive feedback, as well. Teachers felt that it was impossible to include all the elements of a standard in one lesson. It was also difficult for teachers to accept that they would not receive the top rating of “6” on all of the sub-standards. Some noted an increase in “drop-in” classroom visits, which concerned some. One individual referred to them as “stealth” evaluations. Principals in Juniper found the pilot to be more time-consuming than the existing system, but also more collaborative. They thought the new system provided more opportunities for professional growth, but some also described it as “cumbersome.”

Labor-management relations reach a new low during an arduous round of negotiations in the spring of 2014, at the conclusion of the first pilot year. Still, the district and union were able to come to agreement that the evaluation pilot would become permanent as of the 2015-16 school year. The agreement states that a new teacher evaluation committee – the third since this work began – would be formed to review and revise the pilot evaluation. The committee structure was explicitly defined. There would be four teachers, including a secondary teacher, elementary teacher, SPED teacher. There would also be three district administrators and one board observer who would not have a vote. Previously, school board members had not been invited to committee meetings. Consensus would require the support of five members. Since changes would need to be approved by a margin of five to three, teachers would have to have the support of at least one district administrator in order for something to pass. As the first year of the pilot concluded, it seemed that the union leadership was not as concerned about the new teacher evaluation system as they had previously been. This may be due to the fact that the stakes were not really high and also because they were able to put another committee in place that would be charged with making improvements.

Stage 6: Continue Implementation

The pilot entered its second year in the fall of 2014. Permanent teachers who participated in the first year would not be evaluated again, because they are either on a two-year or five-year evaluation cycle. The third teacher evaluation committee met once in the fall and was scheduled to meet three times in the spring of 2015. The committee nominated the Assistant Superintendent of HR to serve as the facilitator. While the superintendent and union president did not sit on the committee, they were tasked with reviewing meeting minutes. Only one individual – a union representative – served on all three committees. He is the lead representative for the union on the current committee, officially serving as the co-chair. Stakeholders hope that having the committee’s structure explicitly laid out in the contract would help move their work forward. The committee will make any recommendations for policy changes in the spring, before the new system becomes permanent

New administrator evaluations were also being rolled-out in 2014. The district continues to work with principals on professional development and coaching, as part of the teacher evaluation system. They are promoting the idea of being a “coaching principal,” making sure that administrators are regularly in classrooms and engaging with teachers.

Stakeholder Relationships

Juniper’s labor-management interactions began from a civil state, with limited trust, and varied between civil and contentious over the course of the reform process. The new superintendent took office after discussions about teacher evaluations began and quickly joined the effort. She formed many joint district-union committees in an effort to foster collaboration. While the district was able to move forward with the teacher evaluation committee, there was not universal buy-in. Perhaps due to the limited trust between labor and management, concerns about how the district intended to use a new teacher evaluation system persisted and posed a continuous challenge.

In the first teacher evaluation committee, interactions were civil and somewhat cooperative, and the relationship between the superintendent and union president at the time was collegial. They were both committed to working together. When new union leadership was elected and the second committee was formed, that relationship was essentially reset and communication became difficult. The union’s response shifted. This second committee had a more challenging task than the first: defining and discussing policy options. Labor-management interactions, both within the committee and outside, were increasingly contentious. At the same time, contract negotiations were particularly adversarial, sparking union protests. These interactions align with the “solidarity” action frame described by Gates and Vesneske (2011). In this approach, union leaders may mobilize their membership and use their support to persuade the district to change positions. They note that, while this strategy can be effective at moving negotiations forward, it can have a damaging effect on relationships, and that was observed in Juniper. Though the committee in Juniper came to an agreement on a pilot evaluation, there was a lack of enthusiasm about moving forward. The union did not fully endorse the effort, which they saw as a district initiative, rather than one driven by mutual interest.

When the third union president was elected and the new Assistant Superintendent of HR joined the district, relationships did not improve. Labor-management conflict at some school sites caused implementation challenges, and the underlying current of distrust persisted. Miscommunication with regard to implementation exacerbated the situation. Many teachers still believed that the new teacher evaluation system was designed to be punitive. Those persistent suspicions made the parties less likely to be cooperative (Walton and McKersie, 1991), and labor-management relations reached a low point. After the first year of the pilot, there was less conflict and contract negotiations went more smoothly. The district and union agreed to form a third committee to evaluate the pilot and there were signs that the labor-management relationship

was improving. There was a shared feeling that protracted conflict was harmful to district morale. Still, as of the beginning of the second year of implementation, the relationship would be best described as civil.

Appendix F. Sumac Unified School District Case Summary

Stage 1: Consider Change

Motivation for Reform

In 2010, when one school in Sumac received a federal grant and successfully drafted an MOU to change their teacher evaluation process, district stakeholders realized that, in order to take advantage of similar funding opportunities in the future, they would need to change their approach to teacher evaluation. While the decision to pursue these reforms came from within the district – there were no outside mandates – teachers did pay attention to the national conversation around student test scores that seemed to be “accusing teachers.”

In 2010, prior to the arrival of the current superintendent, the issue of teacher evaluations came up during IBB contract negotiations and the idea of forming a committee was suggested as an option. The current union president had been involved with CTA’s teacher evaluation work at the state level and was also interested in working on the district’s evaluations. As in other districts, teachers were especially concerned when the *Los Angeles Times* released the value-added scores for teachers in the LAUSD. They also paid attention to conversations about merit pay and legislative attempts to limit collective bargaining in other areas of the country. Since the district’s PAR program was taken away during budget cuts, teachers have needed more support, especially with the new Common Core standards. Stakeholders agreed that there was a lack of consistency in the existing teacher evaluations and that it was not a meaningful or helpful process. One stakeholder commented on the challenge of implementing Common Core and that an effective evaluation system should help teachers be successful:

How do we create better teachers to meet this new rigor for Common Core? How do we talk about our areas of expertise and then our areas of need without reflecting upon what were good at and what we need to work on? So, that's where the process started.

Further, many other districts in the region were updating their evaluation systems as well. The hope was that an improved evaluation system would involve more conversation and be better linked to support for teachers. At the same time, Sumac was subject to a lot of outside criticism, which may have motivated efforts to implement reforms. As one individual noted:

So personally I believe a lot of the issues that we've had, as it relates to student achievement and student achievement and our overall achievement as a school district, can be attributed to the fact that there was never really a strong evaluation system. And I don't mean that's because we have bad teachers, I just mean it's because that dialogue was never taking place.

It seemed like the right time to tackle teacher evaluations.

Decision to Consider Change

In 2010, under the previous superintendent, the district formed an exploratory teacher evaluation committee chaired by the existing union president and former Assistant Superintendent of HR. However, there were no clear outcomes of that committee's work. The current superintendent came to the district in early 2011, just as the district was finishing contract negotiations. She communicated her goal of having effective teacher evaluation system, having worked on teacher evaluations in her previous district. She also expressed that she wanted to be a "team player."

The new superintendent identified three characteristics of the existing system that needed to be addressed: 1) principals were not in classrooms long enough, 2) the evaluation did not require lesson planning, and 3) they were not using student data to drive instruction. In fact, when they looked at data, some students were moving backwards. Some stakeholders remained a skeptical of the new superintendent's goals. The superintendent proposed the idea of re-starting the teacher evaluation committee at her monthly standing meeting with the union as a way to address their concerns. Union leadership agreed to move forward with the committee.

Committee Recruitment

The process of forming the committee was quite collaborative. The union and district leadership sat down together to discuss the ideal committee composition. They wanted to make sure teachers from every grade level were represented and that the group included district administrators from different departments (e.g. human resources and educational services). They were very deliberate in who they chose as their representatives. The newly formed teacher evaluation committee would meet for the first time in the fall of 2011.

Stage 2: Set Goals

Sumac's teacher evaluation committee consisted of both district and union leaders from different departments, as well as representatives from all grade levels. They spent the entire 2011-2012 school year in the goal-setting stage. When the committee began meeting in the fall of 2011, the superintendent shared her goals for the process; she wanted to have an effective process to help teachers, which would, in turn, help students. According to committee meeting minutes the superintendent "urged the committee to work collaboratively to develop a meaningful, non-'gotcha' evaluation system for certificated employees."

The committee started by developing a shared understanding of "effective teaching" and coming to an agreement on the purpose of teacher evaluations. They looked at a number of resources on teacher evaluation. They first spent time with a tool from the NEA's resource guide on teacher evaluations called, "Moving to a Shared Understanding of Effective Teaching" (National Education Association, 2011). They used that tool to reflect on what they envisioned an evaluation system to be, before they began looking at different evaluation models. They

identified the three most important instructional practices, influences, outcomes, and areas of knowledge, which would all support effective teaching. They completed these thought exercises individually, at first, and then shared with a small group before engaging in a whole-group discussion. They decided that the purpose of an evaluation was “to strengthen the knowledge, skills, and practices of educators to improve student learning.” Additionally, the group explicitly noted that making that purpose clear to all stakeholders would be essential when soliciting buy-in, especially later during the implementation phase.

They used an activity outlined in an NEA guide on developing teacher evaluations, called “Moving to a Shared Understanding of Effective Teaching” (NEA, 2011). Following this tool, committee members first reflected individually, then came together for a group discussion on their vision of effective teaching and the purpose of an evaluation system. They agreed that the purpose of a teacher evaluation should be “to strengthen the knowledge, skills, and practices of educators to improve student learning.” One individual summarized the guiding principles they agreed to as:

Basically, that it's a team effort that both sides have input, that it's collaborative, not a gotcha. It's confidential obviously. The whole idea is to promote a supportive climate between both sides. They want it to be transparent. That's another big word they used a lot. Those would be some of the main things.

Sumac's committee did not spend much time on research at this stage. They did attend a speech by Robert Marzano about teacher evaluation and begin compiling a binder with resources, but they did not have tasks to complete outside of the meetings. The group decided that it was important to first develop guiding principles of their own before looking at examples from other districts. As one individual recalled:

And it was very clear, "No, we're not doing that. We're not going to go out and look at what other people do. What we're going to do is something different." And so we sat down and did as we said, "Let's first begin with a discussion about what an evaluation system should be." And based on that, we developed some very clear guiding principles.

Some stakeholders described the work they did in the goal-setting stage as the most difficult part of their reform efforts. They agreed on the mission of their committee and developed guiding principles to structure their conversations about evaluation systems. The mission was to:

Develop, implement, and maintain an evaluation system that strengthens the knowledge, skills, and practices of educators to improve student learning.

Their guiding principles reflected the consensus that evaluations should be standards-based, differentiated to match various teaching assignments, and include multiple sources of evidence, both formative and summative. They noted that the evaluation process, as a system of professional growth, should actually promote collaboration between teachers and administrators. Another guiding principle was that the system they adopt should be research-based and must have “clearly defined components and procedures.” The committee agreed that there should be

joint-trainings on the new system to teachers and administrators, rather than introducing the new policy to them in separate forums.

Sumac's committee did have some difficult conversations. They actually discussed the accountability movement as it was playing out in the *Los Angeles Times*' series of articles on value-added measures of teacher effectiveness. One participant described the committee's conversations on the topic:

So we talked about how test scores were published and teachers' names were published, and was that a good thing? Did that advance things or did that hurt things? And so, most of us in the room I think agreed that standardized tests and assessment are a part of what a teacher can use, but that putting it specifically in an evaluation doesn't serve the purpose that we want it to.

Some committee members were surprised that the district and union representatives were on the same page about so many things. The group agreed that they wanted to change the evaluation process without being punitive, and that they wanted to cultivate a coaching environment, rather than a "gotcha" environment.

The committee created a webpage to communicate their work with the rest of the district, including the school board and all the school sites, and also maintained a dedicated email address to solicit stakeholder input; those emails were automatically forwarded to all committee members. After every meeting, they sent an update to the district, in an effort to maintain transparency and provide the community with the opportunity to offer feedback. All of the committee's notes were stored on the website and were accessible to anyone who wanted to provide input. They also encouraged stakeholders to have their own conversations about teacher evaluation at each school site. As one committee member recalled:

We had conversation topics for about a year, about various aspects around the standards and around teacher evaluation that we asked every school site to engage in.

During this time, union representatives went to school sites to listen to members' concerns.

The goal-setting stage took a full school-year, but was considered to be very important. One stakeholder actually described the process of developing guiding principles as "mind-numbing:"

What are our guiding principles? What are the things we believe? That, to me, was mind-numbing to get to that place. I would say it was no less than, if we say, six-hour days times like five and like 30 hours on that. Just unbelievable. I thought my head was going to explode.

However, that same individual recognized how crucial that painstaking process was to the success of the evaluation committee:

You have to build relationships and when you take the time to build the trust and relationships then, you can get the job done faster. So, I've learned to very much appreciate that, but I used to kind of dread the meetings.

Most individuals agreed that the lengthy and challenging goal-setting process was worth the effort.

Stage 3: Define and Discuss Options

After investing a significant amount of time into building trust in the goal-setting stage, Sumac was ready to evaluate policy options. At the beginning of the 2012-2013 school year, there was a joint meeting with the union and district representatives. They shared a PowerPoint presentation to orient members and tried to “grow the group.” As many as 22 individuals participated in Sumac’s committee at this stage, though not all at once. The committee met monthly throughout the school year, and essentially spent that time educating themselves. The group reviewed a lot of literature, including works by Linda Darling-Hammond, Jim Popham, and Robert Marzano. They looked at different policies, including the work of neighboring school districts, and compared them to their existing policies. They described that work as important:

We were being critical of what other districts were doing and supportive of other things. And little by little, what began to transpire is, “Let’s take a look at what some of these other places are doing. We don’t have to re-invent the wheel. Most of the districts around us, all the districts around us had already shifted to the CSTPs to begin with. So let’s go down that road.”

The group paid attention to what implementation might look like their own district and how they would be able to ensure successful implementation. One committee member described those conversations:

“Why is this the case? Why is this working? Why is it not working? Why would this not work in [our district]? Why would it work in [our district]?” And then once we agree on something, “What are we going to have a do to make sure that it works?”

Their work was purposeful and they did not rush; they continued to “go slow to go fast,” as one stakeholder described:

When you consider two and a half years, a group of people invested two and a half years to get to this point, and we read that research, “Go slow to go fast. Go slow to go fast.” Let’s not rush. Let’s question. And I think that’s the other thing that promoted the discussion.

Members said they continued to be treated as equals and that interactions were friendly and respectful. They decided to create sub-committees to put the pilot together, which they found to be an efficient approach. When the current Assistant Superintendent of HR joined the district, they were already in the middle of debating policy options. He took on the responsibility of creating different forms for every certificated position, based on the standards of that profession. He then sent out the forms to get feedback and make sure they captured the essential elements of each position. The superintendent remained behind-the-scenes during these stages, though she kept abreast of the committee’s work.

Throughout these stages, Sumac's committee put effort into engaging the larger district community. The district put together a video and principals were required to show it at faculty meetings, so that everyone was getting information about the work being done around teacher evaluation. They released a list of topics around teacher evaluations that they expected schools to discuss. The committee made an effort to be transparent and to solicit feedback. The district met with principals. The union leaders were proactive and met with their elected representatives. They also went to every school site and had one-on-one critical conversations with members, so that they could best represent them. The committee maintained its website and email address and created an online survey about teacher evaluation, which was open to all teachers and administrators. The committee also reported to the school board at open meetings to keep them updated. However, the school board members were not involved in the committee directly and did not try to micromanage the committee's work. They trusted the leaders of the district and the union. There were no major public comments about teacher evaluations at school board meetings.

Stage 4: Come to Agreement

The committee operated under a consensus mode; all decisions had to be unanimous. After a full year, they finally came to an agreement. They accomplished this outside of formal contract negotiations, which was significant. Early on, the committee made a tacit agreement that value-added measures were off the table and that test scores were not going to constitute a certain percent of teacher effectiveness.

Though the possibility of using value-added measures was taken off the table, teachers had lingering concerns about student achievement data. In Sumac, the district would have liked for student achievement data to play a role in teacher evaluation, but knew that would be a deal-breaker:

For the eval, the CSTP has guided pretty much the whole criteria. There was a... I don't even know if it was really a big debate or not, 'cause we kind of already knew it was a foregone conclusion. We really did want a very specific data piece in there, and we knew the teachers wouldn't be comfortable with that given how student data has been used against teachers in a pretty unfair fashion in other districts. I think it was a year ago or a year before that LAUSD had all those teachers names get published in the Times and all their student achievement data was there.

However, Sumac came closer to including student achievement data as a measure of teacher performance than the other case study districts, by including the following sub-standard:

Demonstrating student progress towards the attainment of grade-level academic standards as evidenced by results from multiple performance measures.

In order to receive the highest rating on that element, teachers must show evidence of "uniform student progress." Further, some stakeholders hoped that student achievement data would be

incorporated into teachers' goals for each academic year. One individual provided some examples of what that might look like:

One of those goals, if not both, should have data components to them. I teach ELs and I really want to work on moving my kids up on the CELDT, or I teach AP English and I really want work on my passage rate for my seniors. I teach Geometry, and I really want work on how the kids are doing on our common assessments. The teachers can come up with those because they do need to gauge, somehow, the data and how they're doing. It's not the be-all, end-all. It's not everything, but they were comfortable with that, as long as the teacher had input as to how the data was going to be looked at or what sort of data.

This was set as an expectation, rather than a formal requirement, and reflected the value the district placed on student achievement data.

Sumac's committee decided what they wanted was to have a coaching practice around the CSTPs. While student achievement is referenced in all of the CSTPs, it was not considered as an evaluative measure for teachers. Their CBA actually prohibits the use of test scores in evaluations. Conversations about student test scores were difficult, but they did come to a compromise that test scores would be involved as part of a goal-setting process for teachers, to be done with the principal. Additionally, Standard 5 says that teachers should use multiple sources to assess student learning. Stakeholders noted that teachers already did a lot of data analysis at their school sites and in Professional Learning Communities (PLCs), and they expected that work to continue.

Sumac's committee discussed the number of levels they should use. They came to the agreement that each sub-standard would be rated on a four-point scale: (4) exceeds standard; (3) meets standard; (2) partially meets standard; or (1) does not meet standard. Using those ratings, a teacher would receive an overall rating for each of the six umbrella standards. The criteria for determining the overall ratings are as follows:

- Exceeds standard: At least three 4's and no 2's or 1's on the related sub-standards.
- Meets standard: A maximum of one 2 and no 1's on the related sub-standards.
- Partially meets standard: Two 2's or one 1 on the related sub-standards.
- Does not meet standard: Two or more 1's on the related sub-standards.

An assistance plan would be required if any overall rating of '2' on any standard was received. A teacher who received an overall rating of '1' would also be required to participate in the district's PAR program. Peer review was only discussed as something that could be formative, but would not contribute to a summative assessment of teacher practice. Those teachers would have to be evaluated annually until they were able to meet the standards.

One of the most difficult conversations was around how long classroom observations should be. Principals normally do walkthroughs and are expected to be in classrooms regularly, but the committee wanted to keep those separate from the formal evaluation process. This was described as an "elephant in the room" until it was addressed. Keeping principals' routine walkthroughs

separate from formal evaluations was considered important. It made the process more collaborative. One stakeholder described why this was a sensitive topic:

Walkthroughs were a main point of contention in this district, because they didn't want principals to evaluate on walkthroughs. You can't help but keep it in the back of your mind, but contractually, we're not supposed to evaluate people on what you see in the walkthroughs, but it is a component. Maybe not data kept, but it ends up... In the new model, you can utilize it by giving an overall picture but not calling it a walkthrough.

The committee decided that the informal observations would continue and could be used to trigger a follow-up observation if something is noticed, but teachers will always know if they are being observed for the purpose of the formal evaluation. There was also a lot of debate over the element of Standard 6, rated to “community conversations.” Teachers were wary of being evaluated on things outside of their classroom. There still are some misconceptions about and trepidation around this standard.

Sumac was able to begin implementation of the pilot without going through formal contract negotiations. The committee came to a mutual agreement on the terms of the pilot and decided that it cohered with the existing contract language. As one individual explained:

So, it's more of a gentleman's agreement that we did through the committee rather than do a formalized MOU. Because the way we saw it, and it was probably a selling point really, was that we're not really veering that far away from contract language. We're veering away from what the system is based on, CSTPs...

Their ability to agree to a pilot outside of negotiation speaks to the collaborative labor-management relationship they were building.

Stage 5: Begin Implementation

Sumac held a joint leadership summit in the summer of 2013, which helped to get district and union leaders on the same page and ready to implement the teacher evaluation pilot. It was described as a success. Before the pilot began, the teacher evaluation committee invited principals and union site representatives from the five pilot schools to one of their meetings. The union and district led training to get administrators calibrated on the rubric. They made a deliberate effort to train teachers and administrators at the same time, so that everyone was getting the same information. As one stakeholder noted:

We're sharing everything with everyone at the same time. This is, we don't want there to be any hidden ideas, we don't want you to hear one message here and here a different message.

Though only five schools participated in the first year, district leaders wanted all teachers to be aware of the pilot. The teacher evaluation committee created a presentation that a principal and a teacher leader at every school were expected to present to the staff. The committee hosted a

webinar and created a video explaining the pilot that was posted on the committee's website and accessible to all district stakeholders and the public.

While technically still optional, nearly all teachers who were due for an evaluation participated; only one teacher opted-out. In the second year, around three-quarters of Sumac's schools participated. Sumac picked the first five pilot schools strategically, targeting campuses that had good labor-management relations and whose administrators would be open to the pilot. One individual explained their rationale as follows:

And we selected those schools with intentionality, because there was already trust at those sites between the administrative staff and the teaching staff. And we also had administrators in that school that we knew were doing something similar already.

Sumac's teacher evaluation committee continued to meet throughout the first year of the pilot and invited teacher site representatives and school administrators to provide feedback at committee meetings. They also held a larger meeting with all pilot participants in the spring. One committee member described their quality improvement efforts as follows:

And, over the course of this year, while it's been piloting, we've been bringing back the site reps and the principals to meet back with the evaluation committee. And we've been bringing them back to say, "Tell us what's working. Tell us what's not working. Tell us what we think we can do better and what we can do differently." And we've been getting really honest feedback from that whole thing.

In addition to administering surveys, Sumac maintained the teacher evaluation committee's email address, to which stakeholders could send feedback at any time.

Stakeholders in Sumac found the new evaluation to be more meaningful than the existing system and appreciated the conversations they were able to have with their administrators. As one committee member reported:

A lot of the certificated staff have been saying when they come in and talk to the eval committee and report on how it's going, it's the first time in many of their careers that they've had a meaningful conversation with their principals about instruction, and so if they can start talking about that and building up trust and rapport on that level it should impact other aspects of relationship and so far it has, so even at our school that we knew weren't as strong in that area it has been a positive, it has.

One teacher praised the optional self-evaluation that was included in the pilot:

I liked the fact that you could evaluate yourself and you could go into your administrator's office and say, "Okay, this is where I think I am. These are some of the things I think I would like to work on. I'd like you to evaluate me on this standard." If my standard is "classroom management," I might want you to come in at the beginning of my lesson, or towards the end.

Teachers did not have to accept the results of the pilot evaluation if they were dissatisfied with the system and the ratings they received, but one principal said no one had a problem with it.

Teachers in Sumac also echoed some of the same concerns heard in Juniper. There was some concern that not all the indicators or sub-standards made sense for teachers in all grade levels and content areas, such as kindergarten teachers or teacher specialists who do not have their own classrooms. For example, Standard 4.2 says that goals are supposed to be set by teachers and the students, but this may be beyond what a kindergarten student can do. Likewise, it would be difficult to rate a teacher without a designated classroom on “physical environment.” They suggested adding an “N/A” column, which is something that Juniper had built-in to their rubric. Some standards, such as “connecting with the community,” seemed too ambiguous. The committee struggled with how to address this issue. If they provided examples of how to reach that standards, then those examples might be interpreted as required activities or the only means for meeting the standard.

The biggest concern voiced by administrators in Sumac was the time commitment. Finding the time to conduct the classroom observations and associated conferences was challenging for principals, especially in schools where there was only one administrator. Sumac required more teacher-administrator conferences and more classroom observations than Juniper did, with at least two unscheduled observations in addition to the minimum of one scheduled observation. Each observation, scheduled or unscheduled, required a follow-up conference. In Sumac, the scheduled observations could last up to sixty minutes, whereas the unscheduled observations were up to thirty minutes. Sumac required a goal-setting conference with the teacher and administrator in the fall, and a post-observation conference to debrief the observed lesson. Sumac required at least one additional pre-observation conference, for the scheduled observation, and a final evaluation conference at the end of the year. One administrator thought that the number of classroom observations should be reduced:

So it was definitely an issue because, number one, the time to be in the classroom is time. Number two, the time to meet with them to go over the observation is time, the time to type it, unless I can get it all typed up during the observation... It's all time, it all adds up. Plus I have to do everything else that I have to do here, so I never got to the third one, and so that's some, one of the feedback I gave was I think we need to just do two, one scheduled, one unscheduled, unless the teacher is deficient in whatever areas and then we would do the third. That's what I proposed.

An administrator estimated that it took five hours to complete the evaluation for one teacher and noted that he had to evaluate about 18 people during the first year of the pilot. One principal recalled having around nine meetings with each teacher:

But if you do the math, there's nine meetings if you do this absolutely, correctly. Because there's Goal Setting, Pre-Observation, Observation, Post, Drop-In, Post, Drop-in, Post, and then final eval [sic] coaching. So, it's nine meetings, if you're really into them substantively. They're, 35, 45 minutes and a minimum of an hour if you're doing them really well. So nine times, 18 teachers, I think I figured out I had 198 meetings. There's only 185 school days.

The time commitment may be an even greater challenge for high school principals, who have larger teaching staffs. As one administrator commented:

The only drawback has been the time management piece. And at a high school, I've got 108 teachers. So in any given year, there is 60 to 70 [up for evaluation]. It's not equal.

However, that same individual said that, while the process is taxing, he found it to be meaningful:

But, of all the things that I do, this is the most meaningful that I do. It's so... I'm not resentful of the time commitment, but it has been a real jump to hyper-speed compared to what we did before. It's been a real big time commitment.

Another administrator appreciated the opportunity to have conversations with his teachers:

I think as we moved through the pilot, they didn't want this to be contentious, but it really does give you some depth and some rigor, and some nuanced conversation with your exemplary teachers. It also is a tool to truly hold your poor teachers accountable that we've never had before.

Since their conversations focused on the standards, they were able to provide more targeted feedback to teachers. The committee agreed to reduce the number of classroom observations from three to two and combine some of the teacher-administrator conferences.

Since Sumac began implementation without a formal MOU, teacher evaluation policy had to be formally discussed during negotiations in 2014. However, since they had done so much legwork in developing the pilot, the conversation at the bargaining table was relatively brief. The contract language was settled in May, 2014, and the new teacher evaluation system would be implemented district-wide in 2015-2016. Ninety-five percent of the teachers union members approved the contract, and the union had the highest voter turnout that they have ever seen. This success may be attributed to the time the leaders spent listening to members. Still, there are some members who did not pay attention to the work at first and so were surprised when the system was implemented.

Stage 6: Continue Implementation

In 2014-2015, last year of the pilot, the new teacher evaluation system was implemented in 75% of the district's schools, including some with more "problematic" labor-management relationships than the initial five schools. Training is ongoing as Sumac continues to roll-out the system. They are also rolling out a new district PLC program, which may complement some of the professional growth efforts tied into the teacher evaluations. Feedback from schools continues to be mostly positive. The pilot will become the permanent system in 2015-2016. It may be modified again before it heads to the bargaining table in 2015. There is still some concern that the one evaluation rubric does not fit all grade levels or teacher assignments, and the

district is still considering adding a five-year skip for veteran teachers, in order to decrease the administrators' loads.

Today, the teachers union feels like they are treated as equals with district leadership. Stakeholders feel that the school board trusts and supports both union and district leadership. There are still some instances when labor and management may not be on the same page. However, they do try to work together as a team. They cite a shared interest – the best interest of the students. The two leaders of the teachers association meet with the district HR administrators every week. They are trying to implement a coaching model throughout the district, to strengthen relationships and build report. They see this as essential to the success of the teacher evaluation work because, as one individual noted, a teacher is “less likely to follow a leader that they don’t believe believes in them.” They have been doing joint trainings for teachers and administrators to work on those relationships.

Stakeholder Relationships

Prior to engaging in their teacher evaluation reform efforts, and prior to the arrival of their new superintendent, labor-management interactions in Sumac were civil. The labor-management relationship took off on a positive trajectory when the new superintendent, put an explicit emphasis on collaboration, arrived. Some stakeholders were reluctant to consider change, but the district and union were able to come together and agree to move forward. In establishing the teacher evaluation committee, district and union leaders recognized that it was important to have representation of diverse stakeholders.

Sumac spent a year in the goal-setting stage, getting the committee established and building relationships among the members. They devoted significant time to developing their guiding principles, a task that would have been easy to push aside. However, their earlier commitment to that work, to finding common ground from the beginning of the reform process, was essential to ensuring collaboration at later stages. Interactions were described as positive and productive. They did have to resolve some disagreements, but simultaneously were fostering trust. When the committee moved on to define and discuss policy options, they were able to work efficiently and without conflict. They continued to focus on collaboration and engaging stakeholders, releasing joint district-union communications to the community. In some ways, the committee’s approach mirrored the IBB process that the district had used for years. Initial resistance to reform seemed to decrease and the committee’s ability to collaborate was recognized by the district community. As the teacher evaluation work was proving to be successful, the district and union pursued other joint initiatives. The joint-leadership summit held in the summer preceding the pilot provided the opportunity to further strengthen labor-management relationships. Sumac was strategic with implementation, selecting pilot schools with exemplary labor-management relations to serve as models. The committee continued to meet regularly to monitor the pilot and also sustain the relationships they had built. Unlike in Juniper, committee membership was pretty consistent;

many individuals were a part of the process from the beginning. At the end of the first year, negotiations went quickly. While there were disagreements, there was an underlying mutual respect that enabled them to find common ground. The occasional conflict is not a sign of contentious relations, but rather a part of any working relationship. The labor-management interactions at this stage mirrored the “diplomacy” action frame described by Gates and Vesneske (2011), in which leaders are committed to problem-solving, while maintaining collaborative relationships. This approach requires open channels of communication between the district and union and the quick resolution of misunderstandings, which is something Sumac’s leaders made a priority. District and union leaders continue to meet regularly and as implementation continues, they will be monitoring the labor-management climate at school sites. They may encounter challenges as more schools participate, but both sides remain committed to the reform’s success. The district recently received external recognition for serving as a model of labor-management collaboration.

Appendix G. Teacher Evaluation Committees

Table G.1 compares the characteristics of the teacher evaluation committees in each of the three case study districts.

Table G.1 District Teacher Evaluation Committees

| | Laurel | Juniper | Sumac |
|-----------------------------|----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| Timeline of Meetings | 2012 - 2013 | EC #1: 2010 – 2012 EC #2: 2012 – 2013 EC #3: 2014 – 2015 | 2011 - 2015 |
| Meeting Frequency | Once a month | Varied | Once a month |
| Number of Members | Approximately 12 (5 teachers & 4 administrators attended regularly) | EC #1: 8 (5 union, 3 district) EC #2: 8 (4 union, 4 district) EC #3: 8 (4 union, 3 district, 1 school board) | 15 – 22 (Breakdown between district and union unclear.) |
| Facilitators | Union Representative Assistant Superintendent (AS) of HR | EC #1: Superintendent & UP #1/UP #2 EC #2: Superintendent & Union Representative EC #3: AS of HR | Previous AS of HR Current AS of HR Union Representative Union President |
| Participants | Union Representatives, Teachers, Superintendent AS of HR, District Admin, School Admin | Union Presidents, Union Representatives, AS of HR, District Admin, School Admin, School Board Member* | Union Representatives, Superintendent, District Admin, School Admin |
| Final Products | Letter of Recommendations, Template of example evaluations | EC #1: Evaluation rubric and forms/documents. EC #2: Revised evaluation forms, began pilot | New evaluation system and pilot |
| Compensated? | Yes | Yes | No, only coverage for missing class |
| Consensus/voting | N/A (Did not have to vote on anything) | EC #3: 5 out of 8 to approve | Consensus – adopted what they use in IBB |

*Note: The school board member was added to Juniper's third committee as an "observer."

As previously noted, leaders thought that representation from different stakeholder groups was important to moving the effort forward. Committees had a fairly even split between district representatives, including school administrators, and union members. The unions recruited teacher representatives from both the elementary and secondary levels. Juniper had the unique situation where the composition of the committee changed twice, in later stages of the reform process. Only two individuals from the first committee remained on the second committee.²⁷

²⁷ The structure of the third committee was prescribed by a Memorandum of Understanding (MOU) developed in contract negotiations. Juniper was the only district that included including a school board member, intended to serve as an "observer," in its third committee.

Evaluation committees generally met once a month. In Laurel, the group met exclusively after school, whereas in Juniper and Sumac, the committees sometimes convened during the school day. Teachers who participated in the committees were compensated for their work in Laurel and Juniper. In Sumac, the district paid for substitute teachers to cover classes if the meetings were scheduled during the school day, but there was no additional compensation provided to teachers.

Committees were jointly facilitated by a district and union representative, with the exception of Juniper's third committee, which chose the current Assistant Superintendent of HR to facilitate. The union facilitator was either the president or a member of the association's board. In Laurel and Sumac, the district facilitator was the Assistant Superintendent of HR. In Juniper, the superintendent assumed that role in the first two committees. The Sumac committee adopted the approach they used in IBB negotiations and made decisions only by consensus. In Juniper, decision-making rules were only explicitly outlined for the third committee, which would need five out of eight members to approve any decision. It was structured that way so that it would be impossible to for a decision to be made without support from at least one representative from the district and the union. The Laurel committee never reached the decision-making stage and therefore did not define the requirements for coming to agreement.

This product is part of the Pardee RAND Graduate School (PRGS) dissertation series. PRGS dissertations are produced by graduate fellows of the Pardee RAND Graduate School, the world's leading producer of Ph.D.s in policy analysis. The dissertation has been supervised; reviewed; and approved by the faculty committee composed of Jennifer McCombs (Chair), Rita Karam, and Julia Koppich.



PARDEE RAND GRADUATE SCHOOL

www.rand.org