

Teachers as Policy Actors: Amplifying Teachers' Knowledge about Teacher Evaluation Policy

Filomena Ethel Hengst

Submitted in partial fulfillment of the  
requirements for the Degree of Doctor of Education in  
Teachers College, Columbia University

2023

© 2023

Filomena Ethel Hengst

All Rights Reserved

## **ABSTRACT**

Teachers as Policy Actors: Amplifying Teachers' Knowledge about Teacher Evaluation Policy

Filomena Ethel Hengst

Despite teachers being centrally implicated in teacher evaluation policy reform, we know little about the effects of its policies and procedures on teachers' work in their school district, and we see limited discussion about the possibility of teachers as policy actors to reflect upon and possibly improve or overcome teacher evaluation. Using data from a teacher inquiry group, informal observations, follow-up interviews, and a document analysis of official documents, the researcher examined the underlying discourses in the official documents of teacher evaluation policy and responses that surfaced when teachers study teacher evaluation policies and practices in their New Jersey school district. The researcher found that federal mandates operated under the premise that improving teacher quality closes the achievement gap, but, in reality, they acted as a punitive measure to target teachers. Additionally, the researcher found the official documents inaccessible to teachers, unresponsive to public comment, and containing standards that are not locally defined. Teachers in the study perceived the enacted teacher evaluation system in this school district to be without much positive impact on professional development. Teachers in the study critiqued the federal teacher evaluation policy mandates for their treatment of teachers and students. They also critiqued the teacher evaluation policies and practices in their school district, recalling episodes of formal observation that inspired negative feelings, administrator bias, and unhelpful feedback. In the end, teachers in the study support each other, remain optimistic and dream up alternatives to make it all work better. This dissertation

concludes with implications for teacher advocacy efforts, professional development, and future research.

## Table of Contents

Prologue .....	1
I - INTRODUCTION OF THE STUDY .....	4
Background .....	4
Problem Statement .....	9
Rationale .....	10
Stance .....	10
Statement of Purpose and Research Questions .....	13
Theoretical Framework .....	14
Significance .....	16
II – REVIEW OF THE LITERATURE .....	17
Teacher Evaluation Reform .....	17
Teacher Evaluation .....	18
Methods of Teacher Evaluation .....	20
Use of Student Growth Measures .....	21
Peer Assistance and Review (PAR) .....	23
Observation .....	24
Student Evaluation of Teachers .....	25
Current Policy Approaches .....	25
Tier 1 .....	27
Tier 2 .....	28
Tier 3 .....	29
The Teacher Evaluation Policy System in New Jersey .....	29
No Child Left Behind .....	30
TEACHNJ Law and New Jersey Administrative Codes .....	31
AchieveNJ .....	33
District-Level Policies and Forms .....	35
New Jersey’s Student Growth Goals .....	39
How Does Reform Affect Teachers? .....	41
Sense of Self .....	42
Autonomy .....	43

Job Satisfaction .....	44
Retention .....	45
Unintended Consequences of Reform .....	45
What Does Teacher Research Reveal About Teacher Evaluation Reform? .....	47
Teacher Research .....	48
Conclusion .....	51
III - METHODOLOGY .....	54
Overview of Study Design .....	55
Researcher Positionality .....	56
Pilot Study .....	58
Research Site and Selection of Participants .....	59
Research Site .....	59
Participant Selection .....	61
Data Collection Methods and Analysis .....	65
Teacher Inquiry Group .....	67
Analysis of Official Documents .....	72
Informal Observations .....	76
Follow-up Interviews .....	77
Methods of Data Analysis .....	78
Limitations .....	78
Presentation of Findings .....	80
IV- UNDERLYING DISCOURSES IN TEACHER EVALUATION POLICY .....	81
Federal Mandates: Theories of Action .....	81
Underlying Discourses .....	83
Shifts in the Conceptions of Great Teaching .....	83
Punitive Logic .....	86
Stakeholder Involvement .....	88
Illusion of Agency at the Local Level .....	94
Conclusion .....	96
V - FINDINGS .....	98
Under Attack and Powerless .....	100
Teachers Under Attack .....	100

Powerless .....	103
Responses to Formal Observations .....	107
Negative Emotional Responses to Formal Observation .....	108
An Unfair Process .....	112
Responses to Other Aspects of the Teacher Evaluation System.....	115
Oversimplified View of Children .....	116
Concerns about the Impact of Standardized Testing .....	119
The Teacher-Administrator Dynamic .....	120
Administrator Bias .....	121
Unhelpful Feedback .....	124
Support, Optimism and Dreams.....	126
Teachers' Suggestions to Improve Feedback .....	126
VI – DISCUSSION AND IMPLICATIONS .....	129
Discussion .....	130
Implications for the Study .....	134
Implications for Professional Development.....	135
Implications for Teacher Research and Policy .....	138
Implications for Teacher Advocacy Work.....	139
Critique of the Study .....	142
Concluding Remarks.....	142
References .....	144
Appendix A: Follow-up Interview Protocol .....	158
Appendix B: Teacher Inquiry Group Sessions .....	160
Appendix C: Verbal Recruitment Script.....	163
Appendix D: Informed Consent Form .....	164

## List of Tables and Figures

		Page
Figure 1	<i>State and Federal Influences on District Teacher Evaluation Policies from 2001-2021</i> .....	30
Table 1	<i>Educator Evaluation Rubric Weights Between 2013 and 2023 Compiled from NJDOE Annual Memos to Superintendents Pursuant to N.J.A.C 6A: 10-4.1(d) and 5.1(d)</i> .....	33
Figure 2	<i>Reproduction of Evolution of Evaluation of Reform in New Jersey (NJDOE)</i> .....	35
Figure 3	<i>The Achievement Gap Between Racial/Ethnic Groups in English Language Arts and Mathematics in New Jersey, 2016</i> .....	38
Table 2	<i>Participant Categories</i> .....	65
Table 3	<i>Data Source – Research Question Correlation</i> .....	66
Figure 4	<i>Teacher Evaluation Policy Over the Years</i> .....	69
Figure 5	<i>Educator Feedback and State Response in Shaping AchieveNJ</i> .....	91
Figure 6	<i>A Photograph of a Consensogram Constructed by Teachers</i> .....	104
Figure 7	<i>Not Just a Number</i> .....	118



## ACKNOWLEDGEMENTS

What a luxury to be able to immerse oneself in their passions! I cannot think of anything better than having space, time, and resources to be able to think critically about the world around me. My world revolves around education, and I am grateful to all of the people who have supported me during my journey studying education at the doctoral level in graduate school.

I am grateful to the participants of the study who came together and with me to discuss issues about teacher evaluation. This study would not be possible without their participation. I see you working hard every day, holding your heads up, and holding each other up, to try make a difference in a system that doesn't value you. The profession has vastly changed since when many of you entered it, yet you persist. You bring your expertise and are always willing to learn something new to educate children. Your commitment to your students inspires me. This dissertation is yours as much as mine.

I would to extend my gratitude to the members of my committee, Dr. Dirck Roosevelt, Dr. Celia Oyler, Dr. Daniel Friedrich, and Dr. Carolyn Riehl. Dirck, I appreciate our conversations. You always pointed out something unexpected, insightful and taught me how to unpack language throughout the process of writing my dissertation. Celia, I would like to thank you for seeing something important in my work and jumping in (and out of retirement) to help me cross the finish line. It was an honor to work with such a brilliant scholar who is committed to liberation, in both scholarship and doctoral advisement. Dr. Friedrich, I appreciate your contributions to my oral defense as well as your lasting commitment to the Department of Curriculum and Teaching. Dr. Riehl, thank you for participating in my dissertation oral defense committee. I would also like to thank Dr. Thomas Hatch and Dr. Karen Zumwalt for their doctoral advisement at the start of my graduate program.

In addition to the scholars mentioned above, I would like to acknowledge the Teachers College graduate students who taught me so much in my journey. I learned as much from them as I did from my professors. There are many more than those named here. I am indebted to my writing partners, Bill Davis and Jennifer Coniff, who helped me tackle the first hurdle of the proposal hearing. I am sending love to my fellow students who made me feel valued during our initiation into the program, Marisa Chin-Calubaquib, Robin Hancock, Wanda Watson, Kay Gordon, Saima Gowani, Rebecca Gomez, Emmanuelle Fincham, Shakealia Finley, and many more. I also have to thank my roommates who acted like a family during a huge time of transition for me as I embarked on this program, Deborah Sheehan, Ann Mullins, Ghazala Mehmood, and Sarah Sweney. I would also like to thank my copyeditors, Lauren and Katy.

My deceased father, Frank Brogna, sacrificed everything to be the father that he never had. He provided us with consistency, loving support and nurtured our resourcefulness so that my brother, Frank, and I could pursue higher education. My father also was the one who put the bug in my ear to become a teacher when I was a teenager. I am proudly the first in my family to graduate college and, now, the first to earn a doctoral degree.

Finally, I would like to thank my husband. Scott, I would not have been able to pursue these goals without your support and the help of your parents, Rich and Judy, throughout the whole journey. I met you serendipitously a few months before I was planning to start my doctoral program, we continued our relationship long distance and decided to wed a few years later. Since that time, we had two incredible children and built the most beautiful live together. Your undying love and partnership have enabled me to achieve my dream.

F.E.H.

*Dedicated to my children, Violet and Daphne*

## Prologue

Throughout my fifteen years of teaching, my female elementary school colleagues and I have often discussed teacher evaluation: the annual formal observations always loom over our heads. Teachers frequently engage in brief conversations to find out where others are in the administrators' formal observation cycles. We congratulate each other if observations go well and reassure each other if they do not.

Thinking back on my formal teacher evaluations throughout the years, from student teaching to present, it is clear that my dream of becoming a licensed practitioner depended greatly on the evaluations of the cooperating teachers and college supervisors during student teaching practicum. Varied evaluations incited mixed feelings in me about my ability as a teacher entering the profession. During my first-year teaching in 2004, I encountered similar mixed messages about my teaching. I felt profound stress that first year, as my job—which afforded me financial stability and a fulfilled promise of my dream to teach—depended on those evaluations. My principal's first observation found me to be generally "proficient" but her second observation deemed me, for the most part, "partially proficient," which alarmed me enough that I thought my contract might not be renewed the following year. Luckily, my third observation saved my job. The vice principal appreciated my creativity in working with another teacher to launch a literature circles teaching strategy. The vice principal became my hero. I felt personally grateful that she "saw" me in the same way I saw myself. The external validation inspired in me hope and relief.

Teacher evaluation was not at the forefront of my mind in the subsequent seven years. I still felt the stress of being observed formally, but evaluations always reflected positively on my

teaching by the end of the process. The job security gained through tenure, along with more collegial relationships with administrators now subdued uneasy feelings I had felt as a first-year teacher in the teacher evaluation cycle.

Changes in teacher evaluation policy at the state and district level, as well as new district leadership during the 2013-2014 school year reignited the stress I felt as a first-year teacher. The collegial relationships I had formed with previous district leadership felt lost due to new supervisors, a new superintendent and a new principal. The observation component of the annual teacher evaluation cycle altered teacher ratings from levels of proficiency to levels of effectiveness. This also symbolized a larger shift that tied teacher ratings to student outcomes locally and statewide. The change created uncertainty around teacher evaluation procedures for administrators and teachers in our district; the change also began a scarcity mindset around ratings, particularly that only a finite amount of “highly effective” ratings could be earned among the district’s teachers.

Interestingly, I did not notice changes in the system and my own ratings until my student teacher, Ceara, pointed out the difference to me in a debrief conversation after meeting with my principal. Ceara’s comments prompted me to compare the rating to previous evaluations. Discovery of the contrast made me question my ability as a teacher and incited feelings of embarrassment in front of the student teacher I was mentoring at the time. To this day, the memory of this evaluation still brings forth feelings of defeat. Other teachers express similar shame surrounding experiences with teacher evaluation since the implementation of the new system. Even though the teacher evaluation cycle seems to come and go quickly, teachers still struggle under the weight of these concerns several times a year. My own fraught relationship with teacher evaluation and the damage to other teachers I have witnessed inspired me to

investigate the issue further. As I set out to conduct dissertation research, I sought to give teachers a platform to share their perspectives on teacher evaluation policy. I wrote central questions to explore underlying discourses in official documents in addition to teacher discourse surrounding the issue. My analysis of teachers' insights from the study led me to conclude that current teacher evaluation policies can be harmful to teachers.

## I - INTRODUCTION OF THE STUDY

The work of female elementary school teachers in New Jersey consists of valuable “knowing-in-action” that could enhance an analysis of teacher evaluation policy. Teacher research is a methodological vehicle to access that knowledge. It also legitimizes the teacher knowledge for consideration by policy, which tends to fall under the neo-positivist paradigm centering a view of knowledge as one verifiable truth. Exploring teachers’ knowledge might also provide insight regarding teacher advocacy efforts. As I set out on the study, I aimed to explore the creation of local knowledge about teacher evaluation by female elementary school teachers engaging in a collaborative teacher research project. Due to various constraints that will be explained later, the hope for collaboration in the research process was not fulfilled. The problem under investigation in this study is what a teacher-conducted, qualitative study can reveal about teacher evaluation policies’ effects on teachers.

### **Background**

Contemporary teacher evaluation policy is premised on the idea that the teacher is the greatest factor in a student success (Anderson et al., 2016), which is a highly contested myth. On August 6, 2012, Governor Chris Christie of New Jersey signed N.J.S.A 18:A, the Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) Act, into law. The TEACHNJ (2012) law operates from a premise that attributes student achievement greatly to the teacher. The various components of contemporary teacher evaluation policy are packaged as an effort to reform tenure decisions for teaching staff<sup>1</sup>. The TEACHNJ (2012) law states that its goal is to “raise student achievement by improving instruction through the adoption of

---

<sup>1</sup> Teaching staff, as defined by the TEACHNJ Act, includes individuals in the positions of teacher, principal, assistant principal, assistant superintendent, school nurses, school athletic trainers, and other employees required to hold appropriate certificates issued by the board of examiners. (New Jersey Department of Education, 2018, p.1)

evaluations that provide specific feedback to educators, inform the provision of aligned professional development, and inform personnel decisions.”

New Jersey’s system utilizes several main performance evaluation methods to arrive at a summative rating of teachers, all of which are tied to the goal of improving student learning outcomes: teacher practice, student growth objective (SGO), and student growth percentile (mSGP). Teacher practice is evaluated by formal observations conducted by a supervisor using a district-chosen observation rubric that includes planning, environment, instruction and professionalism. To produce a formal observation report about teacher practice, many districts in New Jersey decided to use a rubric based on Charlotte Danielson’s (2008) *Enhancing Professional Practice* which was originally developed for learning and professional development. The SGO is a classroom-based objective for the students to achieve, as determined by the teacher and principal. The mSGP is calculated by the NJ Department of Education based on growth in student achievement on standardized tests from year to year. mSGPs are measured by the student performance in New Jersey Student Learning Assessments (NJSLA) in English Language Arts, Mathematics, and Science. The law set forth minimum requirements for evaluation rubrics used to assess teaching staff. For the first time with the passing of this law, teacher ratings became based on students’ test scores for core subject area teachers. For teaching assignments outside of the “core subject areas,” the districts were required to select other methods of measuring student growth. Formal observations conducted by administrators remain a factor in the system, with slight changes to the guidelines pertaining to time requirements and also to encourage districts to utilize various administrators’ perspectives in the process. The law also increases the number of formal observations for non-tenured teachers per year, and increases the number of years before new teachers achieve tenure from three years to four years. The law



emphasizes multiple measures of performance-based teacher evaluation tied to student achievement, and the whole process intends to inform the direction of individual and district-level professional development opportunities. In addition, the TEACHNJ Act intended to inform personnel decisions about tenure.

The 2012 changes to New Jersey's state level teacher evaluation policies grew out of larger federal policy reform initiatives. In 2011, New Jersey won a \$38 million award from the third round of the Obama administration's Race to the Top (RTTT) (Calefati, 2011). RTTT was a competition to encourage and reward states that created plans that focused on college and career readiness standards; student growth; effective teachers and turnaround schools (U.S Department of Education, 2010). RTTT allowed the Obama administration to sponsor the competitive grant program without waiting for the five-year reauthorization of the 1965 Elementary and Secondary Education Act (ESEA), which was known as the "No Child Left Behind Act of 2001" (NCLB). NCLB was the Bush administration's iteration of ESEA, which was remained in limbo at the start of the Obama administration. The ESEA would not be reauthorized until four years later in 2016 under the name "Every Student Succeeds Act" (ESSA). The TEACHNJ Act was created and passed in fulfillment of the objectives set forth by the federal law.

The Obama administration criticized the treatment of teachers under the previous administration's NCLB (U.S. Department of Education, 2010). According to the Obama administration's (2010) *Blueprint for ESEA Reform*, teachers had been unfairly blamed. They suggested that teachers were given cursory evaluations, limited feedback to improve, scattered professional development, no time to collaborate with colleagues, and poor compensation, even when participating in extra duties and when serving in high-need areas. The Obama

administration sought to strengthen teaching with the new evaluation system: it was meant to provide meaningful feedback to teachers through observations of their teaching; teacher portfolios; student and parent data; and student growth data. Under the Obama administration's plan, all teachers should receive support in order to improve student learning. Advanced teachers were to be provided leadership opportunities and other incentives for growth. Principals were to be developed, evaluated, and rewarded if they made way for teachers to thrive. However, Obama's administration utilized the same punitive logic as NCLB, as discussed more fully in Chapter 4.

Both the Obama and previous administrations were influenced by outside groups in constructing these goals. Hursch (2016) identifies the Lynde and Harry Bradley Foundation, the Koch Family Foundation, the John M. Olin Foundation, the Scaife Family Foundation, and the Adolf Coors Foundation as some of the top outside groups. Hursch explains that these wealthy foundations and many others poured millions of dollars into think tanks and lobbying efforts to influence federal mandates since the 1990s. Hursch argues that federal mandates, including teacher evaluation policy mandates, are actually created outside of government by unelected, private individuals who are unaccountable to the public. These groups are fueled by a neoliberal, or market-based, ideology. Teacher evaluation policy mandates were created by outside groups as a way to evade protections afforded to teachers through tenure laws.

The new state-mandated teacher evaluation policy reform went into effect for New Jersey during the 2013-2014 school year. An elementary school teacher of state-tested subjects, such as English Language Arts, Math, and/or Science had three components of their total evaluation score weighted as follows: teacher practice (50%), a SGO (15%), and growth demonstrated by students' performance on state tests in math, language arts, and/or science or mSGP (35%).

Rollout of Common-Core aligned assessments called Performance Assessment of Readiness for College and Career (PARCC) in 2014 temporarily decreased the weight of student test scores in the mSGP aspect of teacher evaluation from 35% to 10%; however, by 2016, the weight of the mSGP component returned to 30%. In 2018, newly elected Governor Phil Murphy and his administration lowered the reliance on student tests in teacher evaluation from 30% to 5% during the 2018-2019 school year. Due to the COVID-19 pandemic, the typical spring standardized tests were canceled for the first time in the history of the New Jersey the during 2020-2021 school year, resulting in teacher practice accounting for 100% percent of a teacher's evaluation.

At the time of this dissertation, as the state continues to face the effects of a worldwide pandemic, an elementary school teacher's evaluation does not contain reference to state assessments and is weighted as follows: formal observations by supervisors or teacher practice (85%), and a self-designed student growth objective or SGO (15%). All components result in an overall, summative rating of the teacher as ineffective, partially effective, effective and highly effective for the year. Despite changes in political leadership at the federal and state level since 2012, New Jersey's teaching staff have not seen major changes to the measures set forth by the TEACHNJ law, even after its official re-adoption ten years later.

Research shows that multiple measures are the best (Darling-Hammond, 2012), and most widely accepted practice in school districts across the country follows this recommendation employing multiple components in their teacher evaluation systems (Anderson et al., 2016), including New Jersey; however, the extent to which those components are used in conjunction with one another can vary widely across districts throughout the country. Many districts use student growth as a component, sometimes with and sometimes without value-added measures, which are the unique contributions a teacher makes to a child's test score absent of outside

factors according to Concoran (2010). Some also use teacher practice instruments, peer models, student evaluations of teachers, and/or other measures of teachers and teaching. Despite different tools, the underlying market-based ideology in teacher evaluation persists.

This narrative describes the recent historical-political context of the most recent teacher evaluation policy in New Jersey based on TEACHNJ. As explained, teacher evaluation in New Jersey is the result of federal mandates that were created through the influence of outside groups with market-based ideology. The system is comprised of multiple measures weighted in accordance with a constantly changing political philosophy regarding the use of students' standardized test scores to evaluate teachers. Contemporary policy around teacher evaluation is premised on the idea that teachers are the greatest factor in a student's success (Department of Education, 2016).

### **Problem Statement**

Berliner (2009) demonstrates that poverty is the greatest factor hindering student success, but that statistic has not overcome the premise in teacher evaluation policy reform efforts that identify the teacher as the most significant variable on the achievement of students (Hanushek, 1992). New teacher evaluation approaches strive to find better measures to dismiss, develop, and sometimes reward teachers (Corcoran, 2010), but regardless of the approach that is used, the teacher is always at the center of the evaluation process with specialized knowledge that is valuable. We know that other forms of educational reform have directly affected teachers' sense of self, morale, and heightened levels of anxiety and dread (Evans, 2000; Troman, 2000) --all of which can have impact on teaching. However, despite teachers being centrally implicated in teacher evaluation policy reform, we know little about the effects of its policies and procedures on teachers' work in their school district, and we see limited discussion about the possibility of teachers as policy actors to reflect upon and possibly improve or overcome teacher evaluation.

## **Rationale**

Teachers' knowledge is valuable, but often ignored. Policy research operates in a neo-positivist paradigm (Ryan, 2015) which positions certain actors as knowers at the top of the hierarchy, whom are usually not teachers. The lack of inclusion of teachers' knowledge in teacher evaluation policy reform is evident. A literature review of contemporary teacher evaluation policy reveals that studies on teacher evaluation reform are not frequently authored by teachers in practice. While it may not be part of a teacher's typical job description to author studies on teacher evaluation or other reform issues, making space for teachers to share their knowledge in this way is crucial to promote equitable and democratic accountability in policy. Federal legislation, such as NCLB and ESSA, strive to provide every child with access to highly qualified or effective teachers, yet, according to Cochran-Smith et al. (2016), fail to address the larger systems that created this inequity in the first place as they are based on "thin" (Barber, 1984) or individualistic notions of democracy. Cochran-Smith et al. (2016) argue for policies based on a "strong" (Barber, 1984), participatory form of democracy. Qualitative research informed by the values of teacher research is a methodological vehicle by which policymakers can take in interpretations, analysis and documentation of teachers' "knowing-in-action" about accountability measures. This provides a formalized way to promote a strong, participatory form of democracy where teachers have an equitable voice regarding policies that impact their work directly.

## **Stance**

University-based researchers and others have challenged the legitimacy of teacher researchers since at least the 1950's (Zeichner & Noffke, 2001 as cited in Cochran-Smith and Donnell, 2006). Qualitative research that enacts or upholds the values found in teacher-led research has the potential to provide insight to teacher evaluation policy. Teacher-researchers are

in a special position as insiders in the professional context to illuminate the distinction between “what is intended and what occurs” in practice, whereas in traditional research, inquiry is based on literature (Cochran-Smith and Donnell, 2006).

Additionally, teacher research can open up the work of teaching and learning to outsiders and potentially inspire social change. Teacher research is outside the traditional epistemological and methodological frameworks which value formal knowledge over “knowing-in-action” and argue that a researcher must be distant and apolitical.

In pursuing this project, I acknowledge my own position as a teacher activist. Sachs (2000, 2001) describes an activist professional’s identity is rooted in democratic principles of equity and social justice. The notion of an activist professional runs counter to the prevailing mode of entrepreneurial professional, or careerist, who responds to social reforms with efficiency and responsibility (Groundwater-Smith & Sachs, 2002).

My development as a teacher activist began in my childhood and has continued throughout my life. I identify as a white, cisgender female who grew up in a third generation Italian American, working class, single-parent household. My father, our main caregiver, was a member of the International Brotherhood of Teamsters, and his union job afforded us stability due to his consistent salary, health benefits, and regular daytime hours. I was drawn to teaching based on his recommendation and in hopes for stability after his death. I struggled to support myself while single earning my teaching salary and had to take on additional work to afford rent within a reasonable proximity. After marrying my husband, who has benefited from generational wealth, we became a two-income, financially stable and home-owning household.

Both patriarchal structures in my male-led upbringing and my proximity to poverty-living in our town’s affordable housing development remain key factors in my conceptualization of my

position to oppression. Due to these formative early experiences, I began to think about systemic change and dream about a different future. High school peers drew me to participate in activism around children's rights and animal rights. In college, I learned about politics and my curiosity led me to participate in electoral politics and support environmental non-profits during my undergraduate years. As an adult and educator, I became active in my teachers' union, local government, and social justice and environmental organizing work in my community. When starting the design of this dissertation, I saw myself as working on social change from within existing policies and regulatory frameworks. This drew me to investigate the policies and procedures behind teacher evaluation rather than pursuing what I characterized early on in my proposal as "some idealistic alternative," (Hengst, 2018).

In addition to an activist frame, my conception of teaching heavily influences my decisions in this project: teaching starts with the purpose of schooling. Cochran-Smith articulates this purpose as "preparing students to engage in satisfying work, function as lifelong learners who can cope with the challenges of a rapidly changing global society, recognize inequalities in their everyday context, and join with others who challenge them" (Goodland 2008, p.xiii). A big part of teachers' work is to sustain American democracy, and the goal of education is to create a more just and decent country (Zeichner, 1991). In light of the country's widespread reckoning with systemic racism and botched response to a global pandemic, I also turn to Bettina Love's conception of modern-day educational abolition (2019) in order to dream about the transformative potential of schools.

As "deliberative democracy asks citizens and officials to justify public policy by giving reasons that can be accepted by those bound by it" (Groundwater-Smith & Sachs, 2002, p. 353), by realizing the stance of a teacher activist, I created a research plan that examined the issue of

teacher evaluation policy in our district with the intention that I was working alongside my colleagues. I also grappled with the idea of organizing our efforts around an attempt to create transformative action in my workplace, based on ideas that came up during the research process. I recognize that my positionality as a teacher and researcher in the school district afforded me more power than the other participants with whom I worked on the study. Since I have gained specialized knowledge about educational research through my doctoral studies, I was seen as an authority by teachers. Regardless, I held their expertise, leadership, and experience in high regard. Working alongside these teachers, having the same teaching responsibilities, and facing similar dilemmas regarding teacher evaluation put me in a position where I was considered an insider and, at times during the data collection, a co-participant. The power relations among myself and the participants were contextual and fluid throughout the study. I was deliberate to “move at the speed of trust” (brown, 2017, p.42) with my colleagues and accept that the will to collaborate beyond discussion did not materialize.

### **Statement of Purpose and Research Questions**

The purpose of study is to utilize teacher research to understand and explain the effects of the teacher evaluation policy on teaching while documenting qualitative methods informed by the values of teacher research to engage teachers as policy actors. This study sought to facilitate a collaborative teacher research project comprised of at least four female K-5 teachers in the same school district where I teach. I engaged in document analysis of official teacher evaluation policy texts to explore underlying discourses. The teachers participated in an inquiry group engaged in discussion about official documents and arts-based, storytelling activities to engage teacher evaluation policy. I analyzed data from the teacher inquiry group sessions in conjunction with data obtained from artifacts, meetings, day-to-day conversations and follow up interviews to analyze the ways the teacher evaluation system is affecting the teaching among this group of



teachers. Qualitative studies informed by teacher research are more successful when there is a university partner (Fregeau and Leier, 2002). Therefore, I participated as a member of the group, and also acted as the lead university-affiliated researcher: I scheduled all meetings, took notes, and facilitated discussion. To examine the effects of the new teacher evaluation policy in the district and document the use of teacher research to engage teachers as policy actors, the study explores the following research questions:

1. What are the underlying discourses in the official documents of teacher evaluation policy?
2. What responses surface when teachers study teacher evaluation policies and practices in their New Jersey school district?<sup>2</sup>

### **Theoretical Framework**

“Our capacity to understand is rooted in our own self-definitions, hence in what we are.

We are fundamentally self-interpreting and self-defining, living always in a cultural environment, inside a ‘web of signification we ourselves have spun. ’There is no outside, detached standpoint from which we gather and present brute data. When we try to understand the cultural world, we are dealing with interpretations and interpretations of interpretations.” (Ranibow & Sullivan, 1987, p. 7)

My approach to this research starts with a social constructivist lens, situated within the interpretivist tradition. As the quotation above illustrates, the research paradigm of interpretivism embraces a complex, interrelated web of meaning in the world rather than one, knowable reality.

---

<sup>2</sup> This question originally referred to the phenomenon as “teacher evaluation reform.” For reasons that became clearer during the course of the study, changes in policies and practices were actually central to the investigation rather than changes signaled by reform. I discuss reforms in Chapter 3 more fully.

Knowledge is socially constructed (Vygotsky, 1962, /1987), and individuals interpret their own meaning of objects and concepts within practical and historical contexts.

In teaching and learning, social constructivism opposes a “banking concept of education” (Freire, 1972). Freire writes, “In the banking concept of education, knowledge is a gift bestowed by those who consider themselves knowledgeable upon those whom they consider to know nothing” (Freire, 1972, p. 72). Rather than positioning only certain actors involved in a school as knowledgeable, such as administrators or policymakers, an interpretivist lens views all actors as competent makers of knowledge.

The theoretical work of John Dewey and that of Donald Schön follow an interpretivist lens. Dewey theorizes that science is the “freed activity of mind” and that all members of the school should have research and reflection embedded in their daily activities without the need for so-called outside experts (Lagemann, 2002, p. 50). Stenhouse (1975) invented the term “teacher-as-researcher” to describe teachers’ commitment to self-questioning and improvement in their practice. These lenses of knowledge and research led to a field of educational inquiry became popular in the U.S. in the 1980s known as “teacher research.” This informs my methodological choice, which I will explore further in the literature review. Schön’s (1983) theory of “knowing-in-action” refers to the specialized knowledge that only practitioners have about their work because of “reflection-in-action,” or teachers reflecting upon their work while doing it. This view honors and legitimizes the knowledge that arises from teachers’ practice and “emancipates people from the positivist “domination of thought” through their own understandings and actions” (Hollingsworth, 1994, p. 8). John Dewey’s democratic research principles and Schön’s

“knowing-in-action” form the epistemological basis of this research as well as honoring a strengths-based perspective of teachers’ knowledge.

### **Significance**

This research project is not just about self-discovery. As the lead researcher and an elementary school teacher of 18 years, I have gathered and analyzed “concrete evidence that the serious and difficult issues involved in teaching, learning, and schooling have been confronted explicitly” (Cochran-Smith and Donnell, 2006, p. 514). Lytle and Cochran-Smith (1992) argue that teacher research can be useful beyond the local context. This research is likely to be of value due to its methodological approach which amplifies teachers’ knowledge in a broader policy context. This teacher research study will be useful to gain insight regarding the extent to which teachers have specialized knowledge to contribute to the broader policy discussion. Given that teachers are considered central to reform, teachers initiating research projects about policy implementation provides a meaningful way for teachers to participate in the rollout of new policy. Additionally, exploration of the effect of teacher evaluation on female elementary school teachers will contribute to the body of research that amplifies female public elementary school teachers’ absent or, possibly, suppressed voices. This research may inform future decisions made by stakeholders in this school district, and New Jersey, regarding teacher evaluation reform. According to Somekh and Zeichner (2009) the influence of neoliberalism on school systems throughout the world (Apple, 2001; Ball et al., 2011) has minimized the role of teachers as technicians. As a practicing elementary school teacher, I seek to broaden the narrow view of teachers’ work and demonstrate how teacher research can explore the political aspects of teachers’ experience that are often overlooked.

## II – REVIEW OF THE LITERATURE

In this review, I take the stance of social constructivist, avoiding rational-technical definitions of knowledge. This is challenging when the dominant discourse in educational policy is neo-positivist or situates knowledge as one truth. By situating the policy in practical and historical contexts, I attempt to resist reifying neo-positivist research that centers the view of knowledge as one verifiable truth and demotes other epistemologies as less valid or reliable. I focus my attention on research and policy that promotes “knowing-in-action” by including studies that try to connect the intended policy aims to teachers’ practice and their lived experiences. “Lived experience acknowledges every aspect of a person's life and identity, even those areas that are not directly connected to the research topic or question” (Given, 2008). In this case, I focus on exploring teachers’ responses and perspectives regarding reform.

Research suggests that educational reform can result in negative changes to one’s sense of self, low morale, anxiety, and dread (Evans, 2000; Troman, 2000). Thus, in an effort to explore the cultural and historically situated reforms in teacher evaluation and its effects on teachers, I focus my literature review on understanding current teacher evaluation reform through policy analysis, exploring the effects of reform on teachers, and considering the use of teacher research in this project.

### **Teacher Evaluation Reform**

First, what is educational reform? Cuban and Tyack (1995) define educational reform as “planned efforts to change schools in order to correct perceived social and educational problems.” (p. 4). One can characterize reform as incremental, which may include small tweaks in new units, decreasing class size, and so on. Or, one can characterize it as fundamental, which may include radical or major changes in desegregation, creation of kindergarten, and so on

(Cuban, 2010). Mort and Cornell (1941) suggest that reforms usually follow these steps: discovering problems, devising remedies, adopting new policies, and bringing about institutional change (as cited in Tyack and Cuban, 1995). People often confuse change with improvement (Cuban, 2010). Even in examples of reform goals that are strongly accomplished as planned, it is still questionable whether there is improvement. These are some complexities that arise when talking about reform.

### ***Teacher Evaluation***

Teacher evaluation systems have become a central topic in contemporary reform since the 1990s. Here, I refer back to Cuban's (2010, April 3) characterization of reform as fundamental or incremental. He says that teacher evaluation systems are an example of a fundamental reform attempt that becomes an incremental reform, since it intends to drastically improve student learning but, in reality, has only made small tweaks in the teachers' experiences with the evaluation process and procedures. For example, No Child Left Behind (NCLB), federal legislation from the late 1990s to early 2000s, focused on ensuring highly qualified teachers taught in every classroom to improve student learning. While teachers' education underwent scrutiny and administrators made personnel decisions based on this approach, students' standardized test scores did not improve. Researchers found that, as years went on, minimum qualifications were not strongly predictive of student outcomes on standardized tests (Goldhaber, 2008; Hanushek & Rivkin, 2006; Kane, Rockoff, & Staiger, 2008).

The Obama administration's *Blueprint for Reform* repeats this narrative that the United States has a failing school system and poor global standing. It says: "America was once the best educated nation in the world. A generation ago, we led all nations in college completion, but today, 10 countries have passed us. It is not that their students are smarter than ours. It is that these countries are being smarter about how to educate their students. And the countries that out-

educate us today will out-compete us tomorrow” (U.S. Department of Education, 2010). This justification was also reported by the current federal legislation Every Student Succeeds Act (ESSA), and used in support of teacher evaluation reform in NCLB and other educational reforms in preceding legislation with The Elementary and Secondary Education Act of 1965 (ESEA). The Reagan administration’s 1983 report *Nation at Risk* and President Johnson’s 1964 legislation *War on Poverty* also made this argument.

Teacher evaluation was under major scrutiny in effort to reauthorize the ESEA when NCLB expired, premised on the belief that the “impact of being assigned to top-performing teachers year after year is enough to significantly narrow achievement gaps” (U.S. Department of Education, 2010, p.13) Obama’s blueprint for reform suggested states/districts should “develop and implement systems of teacher and principal evaluation and support, and to identify effective and highly effective teachers and principals on the basis of student growth and other factors...[to] inform professional development and help teachers and principals improve student learning.” (U.S. Department of Education, 2010, p.4). New policy approaches were adopted to implement the reform, ESEA reauthorization, ESEA waivers, the Race to the Top (RTTT) competition, School Improvement Grant (SIG) program, and the Teacher Incentive Fund (TIF). As a result, differentiated professional development, compensation, and career advancement based on educator evaluation were enacted at state and local levels because of these federal-level legislation, grants, and funds. For example, New Jersey passed the “historic tenure reform” law TEACHNJ in 2013 in response to these funding opportunities and other political factors. The same reforms persisted with the reauthorization of ESEA as the Every Child Succeeds Act (ESSA) in 2016.

### ***Methods of Teacher Evaluation***

In recent years, tracing theoretical and historical influences on teacher evaluation presents a narrative of increasing reliance on student outcomes. No Child Left Behind (NCLB) was focused on eliminating the achievement gap between white students and students of color and establishing the proof that federal dollars are well spent (Noddings, 2007, p. 2). Under NCLB, state tests only measured grade-level standards (Darling-Hammond, 2013), a tradition that continued under the next iteration of that federal legislation. This is an invalid approach, actually, because it does not challenge certain learners who have already mastered grade-level objectives, does not assess growth in students who start out as below grade level, and does not accurately assess students who are English language learners (Darling-Hammond, 2013). Teacher evaluation during the NCLB era was based purely on teacher qualifications, such as actual certification to teach that grade level (Corcoran, 2010), but the growth of invalid standardized tests during that era would lead to other problems for teacher evaluation in the next turn.

In 2009, the Obama administration's \$4 billion Race to the Top (RTTT) competition created an increase in school accountability with a reliance on student achievement on test scores (Corcoran, 2010). The use of student growth data, such as performance on standardized tests, to evaluate teachers increased greatly in recent reforms (NCTQ report, 2013). In promoting RTTT, Obama threatened that states could not ban the use of teacher evaluation that relies on student tests scores, remarking, "Success should be measured by results...That's why any state that makes it unlawful to link student progress to teacher evaluation will have to change its ways" (Corcoran, 2010).

"Linking teacher success to student success or achievement is complex and misleading." (Oakes, Franke, Quartz, & Rogers, 2002, p. 232). Researchers and teachers' unions caution

against using standardized test scores as an indicator for expert teaching. The National Education Association (NEA), the country's largest teachers' union, advocates instead to push for a peer model of evaluation embedded in teacher-led schools (NEA, 2012). However, despite these cautions, teacher evaluation policy today still includes a student growth component that is, when possible, reliant on students' performance on standardized tests. I will explore the use of student test scores further in the next section, as well as other approaches that are currently in use around the country to evaluate teachers.

### *Use of Student Growth Measures*

One example of an increasing reliance on student outcomes is the wider use of student growth measures (SGM), including student growth percentiles (SGPs). SGMs are the approximation to indicators of student outcomes, have grown widely and are featured in most contemporary teacher evaluation systems (Kennedy, 1999). What is the validity of the use of student growth measures in evaluating teachers? Xu et al. (2016) conducted a quantitative study looking at the internal validity of Virginia's teacher evaluation system from 338 teachers in 16 schools, looking at whether teachers with SGPs have different results from teachers without that data, such as teachers who do not teach tested subjects. The study finds that ratings on the process standards such as planning, assessment, and professionalism were actually more predictive of student achievement progress than formal observation of teacher practice. For that reason, the study recommends the continued use of multiple measures for evaluation, including observation, artifacts, teacher self-assessment, student surveys, and value-added assessment.

Corcoran (2010) examined the theoretical notion of value added, which is the unique contribution a teacher makes to a child's progress absent of outside factors. Corcoran argues that if a teacher is going to be assessed by student achievement, then the other factors involving



family, for example, must be factored out. Figuring out a teacher's value-added measure is an important point that must be considered in any system that utilizes student growth data to evaluate teachers.

Still, it is difficult to isolate and calculate a teacher's value-added measure, and stories of its use have been complicated, according to Corcoran (2010), who tells the stories of New York City's and Houston's emerging uses of value-added measures. New York City Teaching Data Initiative sought out to use student test scores for professional development in 2009 and signed on to participate in the large-scale Measures of Teaching Effectiveness (MET) project funded by the Bill & Melinda Gates Foundation to benchmark value-added measures. But, in New York City's application for RTTT, they were required to link student growth data to evaluate teachers and they hastily left out the value-added measures they intended to use, according to Corcoran (2010). Houston Independent School District (ISD) has used value-added measures since 2007 to dismiss teachers and award performance pay: there was even a computer glitch that accidentally overpaid 100 teachers. As a result, Houston ISD overhauled and renamed its program with three strands. One strand addresses school-level value-added measures, the second measures student growth, and the third offers additional performance pay between \$6,600 and \$10,300. Dallas, Denver, Washington, D.C., and Minneapolis also use value-added measures in their teacher evaluation systems. While MET was an impressive and well-funded undertaking, it has faced criticism regarding its conclusions about value-added measures, since students were not randomly assigned to teachers and the influence of high-stakes tests caused teachers to pay special attention to certain content while neglecting non-tested areas of learning (Rothstein & Mathis, 2013).

While the use of student growth measures in teacher evaluation systems became more widespread across states following the Obama administration's RTTT contest in 2009, the prevalence started to decline after the reauthorization of the Elementary and Secondary School Act in 2016. According to National Center for Teacher Quality (2019), in 2009, fifteen states employed measures of student growth. In 2011, that number grew to 23, was at its peak of 43 by 2015, and decreased to 34 states by 2019. Although they adopted the approach at first, Alaska, Arkansas, Kansas, Kentucky, New Mexico, North Carolina, Oklahoma, Wisconsin, and Wyoming retreated from using student growth data in teacher evaluation between 2015-2019. The net gain of states who stuck it out using this approach dwindled down to only 2 states by 2019. Twenty-four states used student growth measures based on performance on standardized tests by 2019.

### ***Peer Assistance and Review (PAR)***

Peer Assistance and Review (PAR) brings together the teachers' union, district administration, and expert teachers, called Consulting Teachers (CTs), to provide support and evaluate teachers using a clearly articulated assessment process (Harvard Graduate School of Education, n.d.). The program is used to evaluate, provide support to, and dismiss new and struggling teachers. The first Peer Assistance and Review (PAR) program started in Toledo, Ohio, in the 1980s as a collaboration between the teachers' union and the board of education, and has persisted for 30 years, becoming the model for other programs all around the country, including at least 41 districts in 13 states (Darling-Hammond, 2012). In Toledo, the majority of late-career teachers who were identified for intervention were dismissed or resigned, with the rest improving their practice because of the support provided (Darling-Hammond, 2012, p. 127). In another district in Poway, California that uses PAR, officials shared that the overall quality of practitioners had been raised by the program. Findings out of Poway suggest that early-career

teacher retention has gone up, with many teachers improving when identified as needing intervention. Teachers who did not improve were dismissed or resign early on (Darling-Hammond, 2012, p. 125). PAR systems offer an alternative evaluation structure to student outcome measures, such as SGPs, which have promoted an increased reliance on student outcomes.

### ***Observation***

Observations of classroom practice are a measure of teacher performance, and the observation cycle may include pre- and post- observation meetings as well as announced or unannounced timing (Danielson, 2013). One requirement for observation instruments is that they must focus on teaching itself so it can be broadly applied to all teachers regardless of content area or grade level. Milanowski (2011) recommends that an observations “go beyond a single administrator with minimal training rating a teacher satisfactory or unsatisfactory based a single observation in a classroom” (p.19). Milanowski recommends that observations should be conducted four to five times to ensure reliability, should include the use of a rubric by a trained administrator, and should be accompanied by work samples and artifacts. One of the most popular instruments of observation in New Jersey is Charlotte Danielson’s *The Framework for Teaching* (2007, 2011, 2013). Danielson’s work also has national significance, as she worked with the Educational Testing Service to develop the Praxis III, which was used by one of the largest studies measuring teacher effectiveness, Bill and Melinda Gates’ large project called Measures of Teaching Effectiveness (MET). Yet, it is important to note again that MET has faced serious critique regarding the validity of its findings (Rothstein & Mathis, 2013).

Other instruments have been approved by the New Jersey Department of Education and studied and analyzed by the state’s largest teachers’ union, the New Jersey Education Association (NJEA). These include the Marshall Rubrics, Mid-Continent Research for Education

and McREL Learning (McREL) Teacher Evaluation Standards, Marzano’s Causal Teacher Evaluation Model, and the Stronge Teacher and Leader Effectiveness Performance System (New Jersey Department of Education, 2015; NEA, 2012).

### ***Student Evaluation of Teachers***

Students may be surveyed to rate the effectiveness of their teachers as another source of data to evaluate teachers. This is an approach frequently used in higher education, and the consequences for faculty vary by institution. Polikoff (2015) points out that the use of student surveys is much less widespread in K-12, so less literature exists regarding the effectiveness in the evaluation of teachers. The results from the MET study show strong correlation from fall to spring and across class sections (Bill and Melinda Gates Foundation, 2010 as cited in Polikoff, 2015); however, Polikoff (2015) notes that there is no evidence about the stability from year to year within the K-12 setting. The MET study presents a hypothesis: student surveys may provide more consistency in the evaluation of teachers than classroom observations.

Carbno (1981) examines the error in students’ evaluation of teacher effectiveness using quantitative modeling. The study points out that there are many variables that are not addressed in the students’ rating. One recommendation suggests prioritizing a larger sample size over a longer list of questions to improve accuracy. Another recommendation is that the students should be identified rather than anonymous to correct for “characteristics known to be related to the ratings” (p. 950), or students’ preconceptions of what good teaching is, for example.

### **Current Policy Approaches**

Policy recommendations saturate the discourse around teacher evaluation. It is helpful to understand them in the practical and historical contexts I describe in the sections thus far. Additionally, before getting into the different tiers of teacher evaluation (Darling-Hammond,

2012), it is useful to consider the theoretical frameworks that undergird contemporary teacher evaluation policy, which are situated in context as well.

Darling-Hammond (1984) points out that the theoretical underpinnings of teacher evaluation are often ignored, but different teaching conceptions are at the foundation of different approaches and, similarly, the relationships between administrators and teachers vary from approach to approach. It is important to consider the portrayal of the teacher when considering varying policies.

Darling-Hammond (1984) outlines different conceptions of teachers and the different consequences in administrative roles for each: (1) The view of *teaching as labor* leads to direct inspection of work, monitoring of lesson plans, classroom performance and a focus on the results. The administration is viewed as the teacher's supervisor. (2) *Teaching as craft* requires repertoire of specialized techniques and their application. A school administrator is a manager who holds teachers to general performance standards. (3) *Teaching as profession* conceptualizes teachers as holding a repertoire of techniques and exercise of judgement (Shavelson & Stern, 1981). The administration provides resources and professional development to ensure competent teaching. (4) The view of *teaching as an art* (Gage, 1978) values personal insight with self-evaluation focused on holistic qualities. The administrator is a leader who encourages teachers. When teaching is viewed as labor or craft, the administrators take a much more evaluative role than teaching as a profession or an art, where the administrator follows the lead of the teacher. The current teacher evaluation approaches are based on a combination of these conceptions of teachers, with less oversight from administrators as a teacher is considered more experienced or advanced.

As mentioned earlier, policy recommendations stress the use of multiple measures to offer more validity (Darling-Hammond, 2012) and to allow teachers who teach non-tested subjects to be evaluated (Anderson et al., 2016). Darling-Hammond (2013) also suggests that a system should consider various aspects of a teacher's work, from instruction to collaboration with colleagues, contributions to one's school community, alignment with statewide standards for effective practice, employment of trained evaluators, and the offering of differentiated professional opportunities or support to teachers based on their evaluations.

Current policies put teacher evaluation on a continuum from initial certification to master teacher status (Darling-Hammond, 2012; NEA, 2012), and there are accompanying approaches for each level. Darling Hammond (2012) describes them in three tiers:

***Tier 1***

In the initial certification process, there are certain standards that novice teachers must meet through performance-based support system, such as the Educative Teacher Performance Assessment (edTPA). The assessment of early teachers is aligned with teacher preparation standards, such as those from the Interstate New Teacher Assessment and Support Consortium (inTASC), National Board for Professional Teaching Standards (NBPTS), and/or the Council for the Accreditation of Educator Preparation (CAEP). Until a few years ago, edTPA was a widespread system used to evaluate candidates looking for initial teacher certification. It was developed at Stanford University through “an iterative process that included 1,000 educators in 29 states and 650 teachers and teacher educators who completed 20+ hours of training to become scorers...According to SCALE (2013), there are currently 655 educator preparation programs in 36 states and the District of Columbia that require candidates to complete the edTPA” (Paine, Beal-Alvarez, & Valdosta, 2016). In the edTPA process, candidates submit segments of teaching to external scorers based on instructional setting and students, three to five lesson plans with

artifacts, videos of instruction, assessments and related data, responses to prompts related to planning, instruction, assessment, and evidence of their feedback to learners. Scorers grade the teachers using a series of rubrics (15-18) focused on teacher performance. Five years after they were introduced, edTPA became even more widespread in the United States (Gitomer, et al., 2021). In 2019, researchers at Rutgers University Gitomer et al. wrote an important article in the American Educational Research Journal (AERJ) challenging the procedures and statistics in edTPA, particularly misleading claims made by the researchers behind edTPA about high reliability of scores and reported low rates of error. In 2022, in the context of this important critique, the COVID-19 worldwide pandemic and historic teacher shortages, both houses of the New Jersey passed A677/S896 to eliminate edTPA a requirement for licensure in the state. There are several other states that eliminated this requirement, including Wisconsin, Georgia, and New York

### ***Tier 2***

The next tier of evaluation is professional licensing (Darling-Hammond, 2012; NEA, 2012), or practicing teachers. The National Council for Accreditation of Teacher Evaluation (NCATE) and, at a state level, the New Jersey Professional Standards for Teachers (NJPST) are examples of standards on which evaluation is based for an early career teacher. Different districts employ multiple measures of performance in different ways according to a case study of eight districts across the United States (Anderson et al., 2016). In Austin Independent School District, the teacher evaluation system is considered a “job appraisal system” with classroom observation, student performance at an individual, team, and school level, student surveys and an administrator’s rating of the teacher on her professional responsibilities. In Washington D.C., the classroom observations, student performance, and principal assessment of professionalism are the contributing factors. In Hamilton County, classroom observation and student performance are

gathered at individual and school levels to evaluate teachers. In Harrison, just classroom observations and indicators of student performance on state and local assessments comprise a teacher's evaluation. The current system of evaluation in New Jersey, the research site, requires formal observations by different supervisors, a self-designed student growth objective, and the growth demonstrated by students' performance on state tests in math, language arts, and/or science (New Jersey Department of Education, 2017). In all of these cases of emerging teacher evaluation systems, multiple measures are utilized to evaluate certified, practicing teachers.

### ***Tier 3***

Advance credentialing and licensing for master teachers is at the end of the continuum of teacher evaluation (Darling-Hammond, 2012). The National Board for Professional Teaching Standards (NBPTS) offers different standards for teachers at this level with implications for teacher evaluation, including an opportunity for teachers to reflect on their teaching. Many indicate that they have grown from this process.

As multiple measures of teacher evaluation are encouraged, common approaches toward teacher evaluation employ varied techniques. However, as stated, all include the use of student growth measures to evaluate teachers, despite questions that arise regarding the complexity and reliability of this approach with a teacher's value-added measure factored into the score.

### **The Teacher Evaluation Policy System in New Jersey**

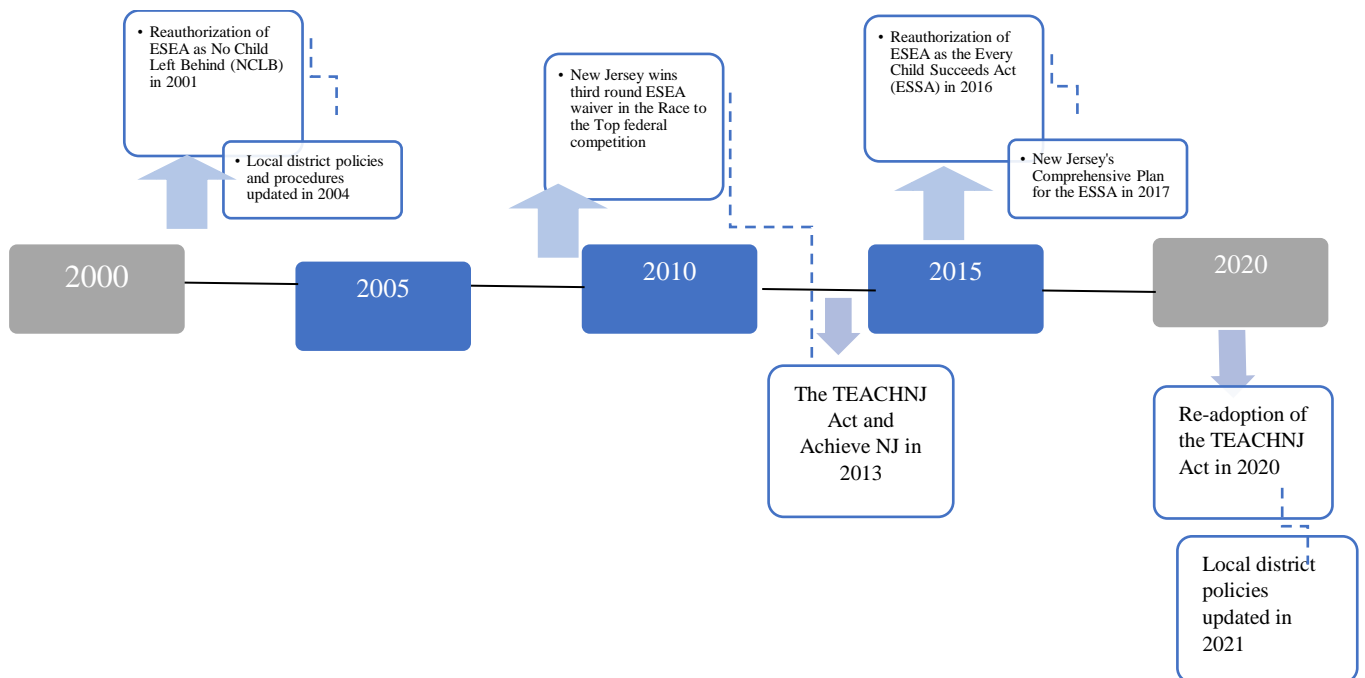
In Figure 1 below is a timeline of the teacher evaluation policy system based on official documents and text from government websites at the district, state, and federal levels between 2001 and 2021. These documents are evidence of the decisions policymakers made regarding the purpose of teacher evaluation and how it should look in each New Jersey school district. These federal influences include the Elementary and Secondary Education Act in 2001 (NCLB), 2016 (ESSA), and the Obama Administration's federal waivers in between. The major state influence



is the Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) Act, which was passed in 2013 and readopted with minor revisions seven years later in 2020. These major pieces of federal and state policy triggered the creation of local district policies in 2004 and 2014, with minor edits in 2021.

**Figure 1**

*State and Federal Influences on District Teacher Evaluation Policies From 2001-2021*



***No Child Left Behind***

No Child Left Behind (NCLB), the reauthorization of the Elementary and Secondary School Act of 1965 (ESEA), was introduced three days after President George W. Bush took office in 2001. On the U.S. Department of Education website ([www.ed.gov](http://www.ed.gov)), he described it as “the cornerstone of my Administration.” It was meant to allow more flexibility for use of federal dollars related to school choice while mandating literacy initiatives and accountability measures for schools that were not making adequate yearly progress. In terms of teachers, the focus was

teacher quality and intended to support the preparation, training, and recruitments of highly qualified teachers. Many teachers recall when districts were forced to send letters to their students' families, stating whether the teacher teaching their child was "highly qualified" or not. The emphasis was on whether the teacher's certification was aligned to the job placement. This federal legislation successfully addressed out-of-field placements in New Jersey and other states, as the number of teachers who were placed in jobs outside of their certification diminished to less than 1% in all subject areas and socioeconomic contexts across New Jersey.

### ***TEACHNJ Law and New Jersey Administrative Codes***

On the website New Jersey Department of Education website ([www.nj.gov](http://www.nj.gov)), the Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) Act was originally passed by the New Jersey state legislature in 2013 as "historic tenure reform." The goal of TEACHNJ was "to raise student achievement by improving instruction through the adoption of evaluations that provide specific feedback to educators, inform the provision of aligned professional development, and inform personnel decisions" (The TEACHNJ Act, 2012). The law introduced a reliance on "student outcomes" to determine teacher effectiveness, which tied student test scores to teacher evaluation for the first time (The TEACHNJ Act, 2012); however, the code stated that it should not be the predominate factor in determining a teacher's annual rating.

In 2013, the passage of the TEACHNJ Act by the legislature and governor triggered an update in New Jersey Administrative Code<sup>3</sup> (N.J.A.C) 6A:10 Educator Effectiveness. Chapter 10, or NJAC 6A:10, stands in compliance with federal rules to provide minimum

---

<sup>3</sup> An administrative regulation or "code" as defined in the Administrative Procedures Act (APA) means "each agency's statement of general applicability and continuing effect that implements or interprets law or policy, or describes the organization, procedure or practice requirements of any agency."

requirements for school districts' evaluation and "to further the development of a professional corps of State educators and to increase student achievement" (N.J.A.C. Chapter 10, 2013). Chapter 10 has nine subchapters that introduce specificity regarding minimum requirements for all district staff evaluation rubrics, use of multiple measures of student performance and measures of teacher practice, guidance for timing around evaluation and feedback, and rules around personnel decisions. As decided by the law, the code specifies that both performance on local assessments and state standardized tests determine a portion of a teacher's evaluation, but the weight of those components has changed several times depending on political climate. TEACHNJ originally required 30% or more of a teacher evaluation to be based on student outcomes and between 50% and 85% based on teacher practice.

When it was readopted in 2020, TEACHNJ lowered the minimum component based on student outcomes to a weight of 20%. Table 1, below, details the changes in rubric weights between 2013 and 2023 for what New Jersey called 1) "mSGP<sup>4</sup> teachers" or teachers of 4<sup>th</sup>-8<sup>th</sup> grade language arts and/or 4<sup>th</sup>-7<sup>th</sup> grade mathematics, where 20 students are enrolled for 60% or more of course duration prior to taking a state assessment. This is approximately 15% of all New Jersey's teachers. The table also shows the weights of 2) those who do not meet those mSGP requirements or "non-mSGP teachers." This group makes up the majority of teachers in New Jersey, approximately 85%. Noteworthy changes occurred in 2016, 2018, and 2020. The rollout of the new Common Core Standards and aligned assessment called the Performance

---

<sup>4</sup> mSGP: a measure of Student Growth Percentiles, which is how much a student improves his or her state test performance from one year to the next compared to students across the state with a similar score history

Assessment of Readiness for College and Career (PARCC) test in 2014 temporarily decreased the weight of student test scores in mSGP from 30% to 10%. Nonetheless, the state felt the initial transition period of getting familiar with PARCC was achieved after three years, so the weight of student test scores rose back up to 30% in 2016. In 2018, Governor Murphy took office and lowered the weight of standardized test scores from 30% to 5%. This persisted for two years until the world was changed by the COVID-19 pandemic. The onset of the pandemic interrupted regular operations of in-person schooling. Standardized tests were canceled during the spring of 2020, and the governor issued Executive Orders 175 and 214 to temporarily eliminate the use of state tests in teacher evaluation. Even though the mSGP measure was eliminated, teachers were still evaluated using locally developed student growth measures and observations of teacher practice. Regular state testing resumed in 2022, and the state reinstated a 5% weight for state assessments in mSGP teacher evaluation.

**Table 1**

*Educator Evaluation Rubric Weights Between 2013 and 2023 Compiled From NJDOE Annual Memos to Superintendents Pursuant to N.J.A.C. 6A:10-4.1(d) and 5.1(d)*

Type of Educator	Multiple Measures of Teacher Evaluation	TEACHNJ Guidelines as of 2013	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	TEACHNJ Guidelines as of Jan. 2020	2021-22	2022-23
mSGP Teachers	• mSGP	≥30%	30%	10%	10%	30%	30%	5%	5%	0%	≥20%	0%	5%
	• SGO		15%	20%	20%	15%	15%	25%	25%	15%		15%	25%
	• Teacher Practice	50-85%	55%	70%	70%	55%	55%	70%	70%	85%	50-85%	85%	70%
Non-mSGP Teachers	• SGO	≥30%	15%	20%	20%	15%	15%	15%	15%	15%	≥30%	15%	15%
	• Teacher Practice	50-85%	85%	80%	80%	85%	85%	85%	85%	85%	50-85%	85%	85%

***AchieveNJ***

AchieveNJ is an implementation support system adopted by the New Jersey State Board of Education to accompany in 2013 the TEACHNJ law. In 2010-2011, the NJ Educator Effectiveness Task Force work started to research ways to reform the teacher evaluation in the

state, and after a pilot program starting in 2011 and new tenure legislation TEACHNJ in 2013, the statewide implementation of a new system called AchieveNJ began. AchieveNJ provides detailed steps for completion of evaluation. Districts can find resources for evaluating teachers and leaders in the form of checklists, scoring guides, and templates to help execute their district's teacher evaluation plan in accord with state guidelines.

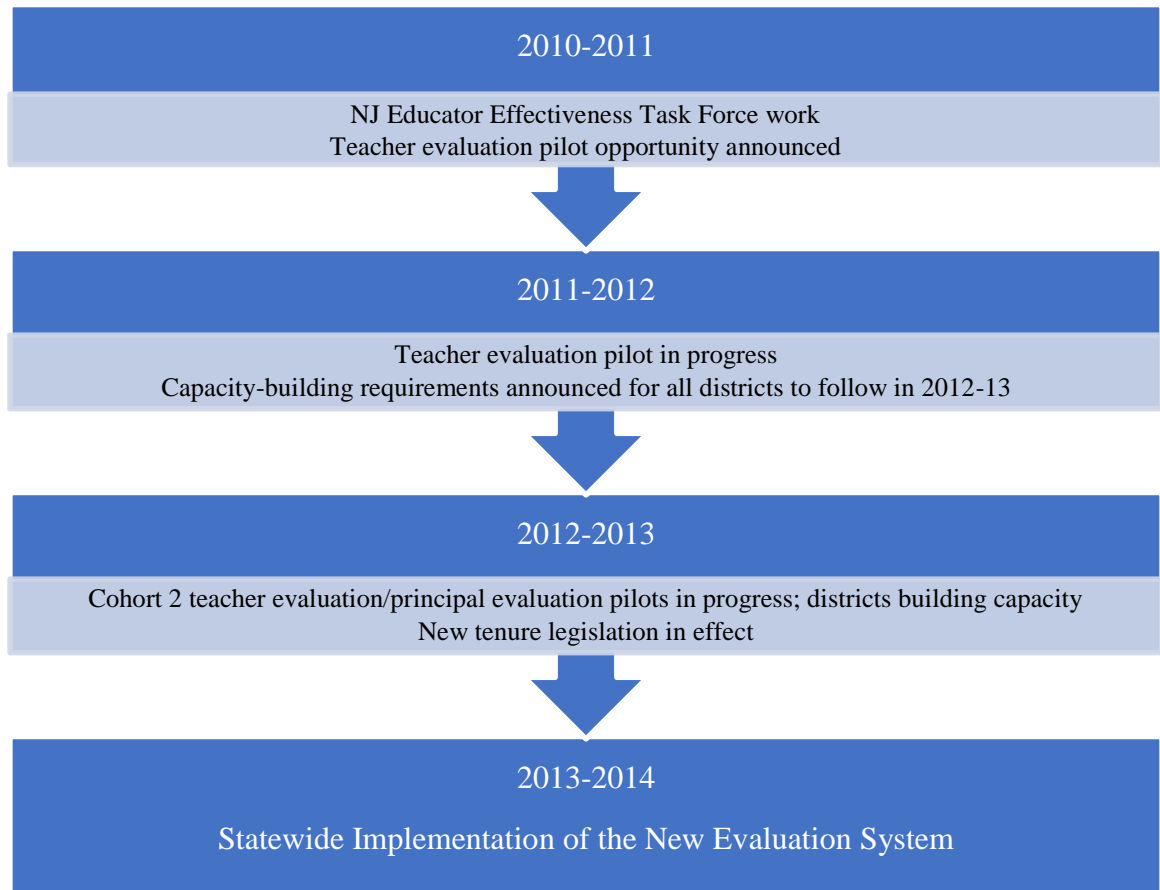
New Jersey piloted AchieveNJ in 30 districts to produce better student outcomes.

According to the state, AchieveNJ “is allowing districts to better identify areas of strength and weakness in educator practice and to respond accordingly” in addition to providing more robust evaluation data for making human capital decisions (New Jersey Department of Education, 2015, p.23). According to the New Jersey Department of Education website ([www.nj.gov](http://www.nj.gov)), the state sees this system as an improvement since it provides “substantially more information about a teacher's performance than was typically available before” and since the teacher's evaluation is based on “multiple measures of performance.”

The guiding principles of AchieveNJ are that 1) the effectiveness of an educator should be measured to ensure teacher quality; 2) evaluations should be based on multiple measures; 3) timely feedback and targeted professional development help educators improve; 4) “evaluation and support systems should be developed with significant input from educators”; and 5) tenure and recognition should be based on effectiveness (New Jersey Department of Education, 2017, October). While New Jersey had teacher evaluation systems before 2010, the rollout of the latest reform is shown in a chart featured on the State of New Jersey's website called “Evolution of Evaluation Reform in New Jersey.” The content and format of that graphic are reproduced below in Figure 2.

**Figure 2**

*Reproduction of Evolution of Evaluation Reform in New Jersey (NJDOE)*



***District-Level Policies and Forms***

Local boards of education updated district policies in 2014 and 2021 to align their policies with new state statutes in TEACHNJ. Districts must be in accordance with New Jersey Administrative Code to receive state and federal funding. The policies in this school district are not original ideas, but they reference state statutes like most of the NJ school districts do for legal reasons. One policy stated that “the Board of Education recognizes the importance of teaching staff member effectiveness to further the development of a professional corps of educators and to increase student achievement,” which echoes the rationale at the state and federal levels for teacher evaluation.

Most school districts in New Jersey use a paid school policy and consultancy agency called Strauss Esmay Associates, LLP to distribute identical policies across state districts that are “kept up-to-date with timely and reliable information based on legal and statutory developments” (www. straussesmay.com). State level rules are reduced to the essentials that districts need to follow to be in compliance with the law. The district policy documents show very little to no original decision-making. The Board policies themselves are simply references to state and federal mandates.

Local district policies give authority in teacher evaluation to the board of education and the superintendent. They outline rules the district must follow in evaluating staff, including deadlines for different components of the system. While the board of education has power to adopt the rubric used to evaluate teachers, such as Charlotte Danielson’s rubric, their role is deferred to the superintendent (NJSBA). The superintendent is the educational expert who administers the school district, including overseeing the teacher and staff evaluation process. Bredeson and Kose (2007) state, “The role of the school superintendent has undergone significant changes over the past century. ... the work has become increasingly defined by the complexities and challenges of political pressure and competing interests (local, state, and nation) ... and greater demand of accountability of increased student performance through state and federal legislation” (as cited in Lazovick, 2020). Yet, superintendents are confined by the political climate in which they operate. As Peterson (2004) points out, NCLB’s rules around teacher quality “make it increasingly difficult for superintendents to lead with a vision of 21<sup>st</sup> century learning,” including parental choice and allocation of resources (as cited in Lazovick, 2020, p.36). Even though teacher evaluation policies seem to anoint superintendents with decision-making power, they feel quite constrained by the mandates.

## **New Jersey's Comprehensive Plan for ESSA**

Federal mandates required that states create a consolidated state plan. In 2017, the New Jersey State Department of Education under Commissioner Kimberly Harrington created one for New Jersey. Starting with the idea that “the vision of NJDOE is for every child in New Jersey, regardless of zip code, to graduate from high school ready for college and career,” (New Jersey Department of Education, 2017), authors of the document also note New Jersey’s remarkable strengths in having some of the highest graduation rates in the country and consistently ranking in the top three on National Assessment of Educational Progress (NAEP) scores. Yet, Commissioner Harrington also frames the policy problem from NJ students’ performance on those same NAEP tests, as well as their performance on the Partnership for Assessment of Readiness for College and Career (PARCC) tests. New Jersey pointed out the disparities among students who are economically disadvantaged versus those who are economically advantaged, as well as students who are minoritized versus those who are white. These disparities provide rationale for New Jersey’s goal-setting under ESSA. The disparity of standardized test scores replicates the logic used at the federal level to justify these reforms. Figure 2 below illustrates these achievement disparities in English language arts and mathematics.<sup>5</sup>

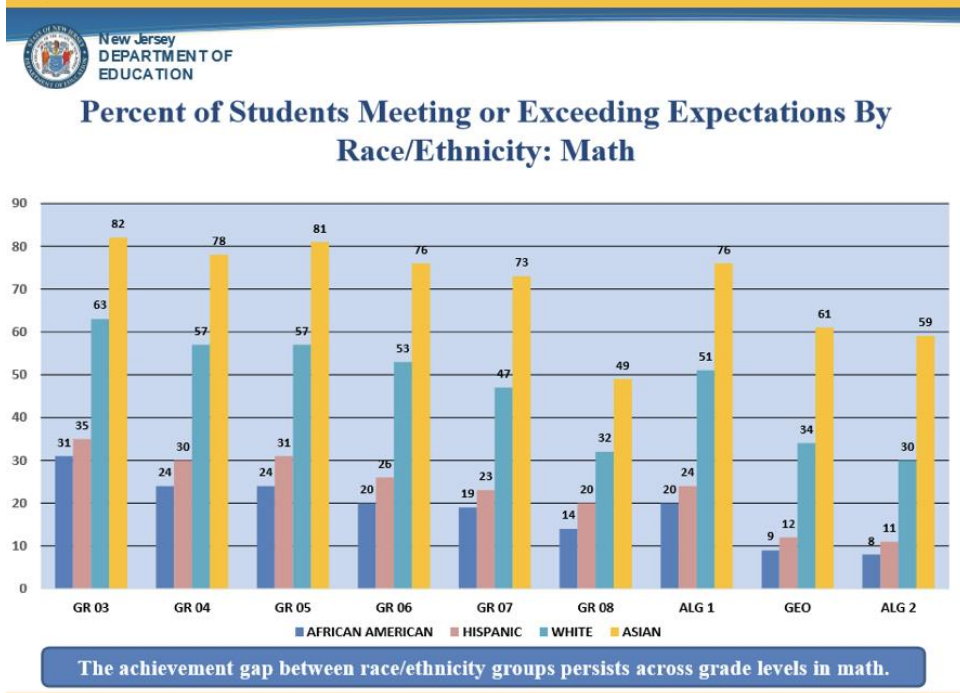
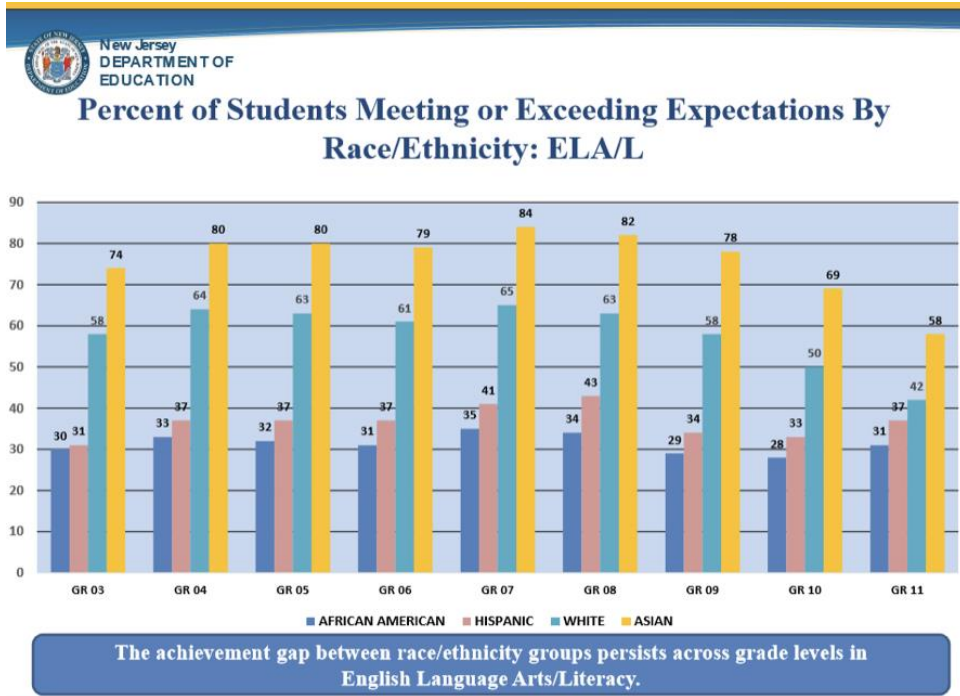
---

<sup>5</sup> The slides are screenshots of the 2016 PARCC results presentation by Peter Shulman and Laura Morana to the New Jersey State Board of Education as provided in Keyes-Maloney, Jennifer (2016, November 3). “State Board Received PARCC Report, Presentation from Newark Superintendent As Board Takes Testimony on Charter School Proposal.” *New Jersey Principals and Supervisors Association*. <https://njpsa.org/state-board-received-parcc-report-presentation-from-newark-superintendent-as-board-takes-testimony-on-charter-school-proposal>



**Figure 3**

*The Achievement Gap Between Racial/Ethnic Groups in English Language Arts and Mathematics in New Jersey, 2016*



In an effort to improve all students' performance on state tests, the state department sought to address a number of factors, including to increase access to what they called "highly effective teachers" for students in low-income settings and students who are minoritized. New Jersey was required by the federal government to collect data regarding the number of inexperienced teachers, those who are novices,<sup>6</sup> and those who are in out-of-field placements. Since No Child Left Behind, there were fewer out-of-field teachers placed in classrooms leading up to this plan, but the high number of novice teachers persisted. According to The IES National Center for Education Statistics, teachers with 10 or fewer years of teaching constituted the 53% of the workforce in 2018 and 50% of new teachers who leave the profession within their first five years (National Center on Education and the Economy, 2007). Research on teacher retention is important but is not the focus of this study.

### ***New Jersey's Student Growth Goals***

In 2017, state testing showed there are significant achievement gaps between "lower- and higher-income students and between minority and white students" (New Jersey Department of Education, 2017, August). The plan also noted the lack of college readiness based on measures like Student Achievement Test (SAT), high school graduation rates, and the number of students taking remedial classes in college.

Similar to how teachers set student growth objectives (SGOs), the state communicated a multi-part growth goal to the federal government. Below are the details of that goal:

- 1) By 2030, 100% of all students will be in either the "approaching," "meeting," or "exceeding" categories on PARCC state testing.

---

<sup>6</sup> defined by the state of New Jersey as teaching fewer than four years

- 2) By 2030, 80% of all students and at least 80% of each subgroup of students in each tested grade will be in the “meeting” or “exceeding” grade-level expectations on the English language arts/Mathematics state tests.
- 3) By 2030, 20% of all students and subgroups will be in the “exceeding expectations” category of PARCC state testing.
- 4) By 2030, 95% of students will graduate in four years and 96% of students will graduate in five years.
- 5) By 2023, 86% of English Language Learners (ELL) will make expected annual progress toward attaining English-language proficiency.

Following the COVID-19 pandemic, New Jersey, like many other states, applied for an extension to meet these goals. In New Jersey’s case, they requested that the 2018-2019 and 2020-2021 school years count consecutively, resulting in an extension of two additional years to meet these goals. Five years after these goals were set, the students’ scores that are used in calculating teacher evaluation scores remain low. For example, as of 2019, just 57% of fourth graders in New Jersey are meeting or exceeding expectations in ELA and 51% of fourth graders are meeting or exceeding expectations in mathematics. Those numbers dipped down to 28% and 26%, respectively, during the COVID-19 pandemic.<sup>7</sup> In another example in 2019, 63% of eighth graders are meeting or exceeding expectations in ELA and only 29% of eighth graders are meeting or exceeding expectations in mathematics. Those numbers also dipped down to 26% and 13% respectively in 2020 during the COVID-19 pandemic. Despite having teacher evaluation

---

<sup>7</sup> These results are from the Start Strong test, which tested 1,000 students instead of approximately 100,000, which is typically tested by the state each year

reform in place since 2013, the state does not appear to be on a trajectory to accomplish its student growth goals regarding state testing.<sup>8</sup>

### **How Does Reform Affect Teachers?**

When I started my dissertation research design, I was focused on how teachers made meaning of reform at the district level. Through various conversations and self-reflection, it became apparent to me that I needed to narrow my topic to one type of reform, such as teacher evaluation. My original literature search offered studies that provided insight into possible negative effects of reform on teachers and suggestions about how to prepare teachers for the changes they will experience in reform, an inevitable part of a teacher's work. I still found these results to be valuable when I prepared to lead a group of fellow teachers into an exploration of teacher evaluation. This results in research from a variety of international and school contexts and highlights the need for attention to teachers' perspectives and emotional responses which I will explore in my study.

While the research that follows in this section shows that there may be negative effects on teachers regarding reform, there are also certain factors that help teachers have the capacity for change. This is helpful to know as we reflect on our varying levels of acceptance or resistance to the new teacher evaluation system in our district. For example, Evans (2000) suggests that teachers who maintain their sense of self and morale may have the capacity for change. This is also the case if they have high levels of metacognition (van den Berg and Schulze, 2014), feel satisfied with their job (Evans, 2000), and are not overcome with nostalgia (Goodson, Moore, & Hargreaves, 2008). This section of the literature review also reveals an insight that policymakers, colleagues, supervisors, and parents have a low trust of teachers, so

---

<sup>8</sup> Additionally, New Jersey's high school graduation rate in 2019 was 91% and in 2020 was 90% which is 4-5 points below its 2030 goal.

working in a trusted context can help teachers better cope with reform (Troman, 2000). The research also suggests that policy could be improved by relying on knowledge from practice more than it currently does (Cohen, Moffit, & Goldin, 2004).

### *Sense of Self*

Teachers' sense of self may be changed through reform unless they already have high levels of metacognition, or awareness of one's thought processes, and are in a context where they are trusted. van den Berg and Schulze (2014) explored teachers' ability to adapt to reforms in South Africa and used the theory of adaptive expertise to interpret the data found in teacher narratives. Amidst the outcomes-based reforms, van den Berg and Schulze found that teachers with high levels of metacognition adapt more readily to changing demands. This shifts the role of the teacher from being an expert who has all of the answers, to a critically thinking individual who facilitates that identity with students as well. Adaptive teachers apply knowledge in new situations and can be innovative when it comes to solving problems.

Van den Berg and Schultz refer to the work of Jansen (2001), who identifies differences in the identities that policymakers construct about teachers versus the teachers' views of themselves. They find that policy images of teachers often seem to exhibit a low trust of their ability to teach. Van den Berg and Schulze (2014) explore teachers' ability to adapt to reforms in South Africa and use the theory of adaptive expertise to interpret the data found in teacher narratives. Amidst the outcomes-based reforms, van den Berg and Schulze find that teachers with high levels of metacognition adapt more readily to changing demands. This shifts the role of the teacher from being an expert who has all of the answers, to a critically thinking individual who facilitates that identity with students as well. Adaptive teachers apply knowledge in new situations and can be innovative when it comes to solving problems.

Not only do policymakers have a low trust of teachers, but as Troman (2000) points out, so do other colleagues, supervisors, and parents. Troman examines the stress U.K. primary teachers face in their school, attributing it to a greater problem “changing trust relations in high modernity” (p. 331). The authors explain that teachers’ work in the last few decades has expanded beyond just classroom teaching to other tasks such as planning, administration, and work with the community. Trust between individuals is a necessary condition for social order (Durkheim, 1956; Elster, 1989) and “greater freedom and autonomy” (Hargreaves, 1998). Troman cites Giddens (1990), who argues that social relationships have become separated from their context, with less likelihood to support basic trust, which requires openness and warmth between two face-to-face individuals. Teachers’ health and well-being are suffering the consequences with a rise of anxiety and dread, the antithesis of trust, in high-risk times of reform.

### ***Autonomy***

Focusing reform on the teacher also challenges a key characteristic of teaching: that it has often been a private act. The lone teacher is autonomous in her classroom, yet this “cellular organization” (Lortie, 1975, p. 23) sometimes leads to inflexibility in adapting reform, a public endeavor. Little (1990) supports Lortie’s findings from many years ago; she argues that “schoolteaching has endured largely as an assemblage of entrepreneurial individuals whose autonomy is grounded in norms of privacy and noninterference and is sustained by the very organization of teaching work” (p. 530). Teaching has been seen as a private endeavor where teachers keep in control in the classroom. Going public can offer praise but can also make teachers vulnerable to criticism. While teachers’ aiding one another or regularly sharing can provide stability and job satisfaction, it is less likely to promote innovation.

### ***Job Satisfaction***

One of the biggest sources of job satisfaction for teachers is found in their work with children. Willingness to change, particularly not being overcome by nostalgia, can help maintain morale during times of reform. Lortie (1975) conducted a sociological study where he surveys teachers regarding their value of three types of rewards: psychic, extrinsic, and ancillary. 76 percent of participants identified psychic rewards, or “subjective valuations,” as psychological benefits they most valued (p. 101) and reported the psychic reward of knowing they had reached children in terms of learning as being career and work rewards of the profession.

Evans (2000) finds that the attitudes for change of U.K. schoolteachers was a significant determinant of their morale in the face of governmental reforms, but individuals’ responses were diverse and personal. Evans looked at a composite study of teachers’ attitudes regarding morale, job satisfaction, and motivation in addition to teacher professionalism, or knowledge, skills, and procedures with which teachers work (Hoyle, 1975 as cited in Evans, 2000), over 5 years. She found that individuals’ responses to change were diverse, reflecting their biographical differences. She noticed that “individuals’ attitudinal responses to change are determined by the extent of compatibility between their own ideologies, values and beliefs and those reflected in the changes they encountered,” as well as the pragmatic assessment of the change’s impact on the system and their own life (p. 185).

Teacher nostalgia may contribute to resistance to change. Nostalgia is more than just an emotional form of memory, it is a “testament of teachers’ experience over of change over time” (Goodson, Moore, & Hargreaves, 2008, p. 43). Goodson et al. argue that change is enacted, not just experienced, by teachers. The researchers found that teachers root their nostalgia in a theme of school as family and feelings of being insulted by reform, which ignores their experience and the complexity of their work. In addition to a resistance to reform, nostalgia

can lead to insecurity about one's capacity, disenfranchisement with students, and pessimism about the future of their school. All of these feelings can contribute to low morale and job satisfaction in times of reform.

### ***Retention***

Having a high level of teacher turnover can lead to poor student achievement and teacher retention can be linked to teacher evaluation reform (Robertson-Kraft & Zhang, 2016).

Robertson-Kraft and Zhang (2016) conducted a quantitative study with interviews regarding the pilot implementation of a new teacher evaluation system in an urban district in Texas. They found that experienced teachers and teachers that were described as "grittier" were most often retained during the reform pilot. Additionally, a small group of teachers were left feeling "burnt out" after all of the additional requirements. Their findings also suggest the value of comprehensive supports (Ingersoll, 2001) and strong leadership from school principals (Boyd et al., 2011).

### **Unintended Consequences of Reform**

Unintended consequences of reform present an interesting paradox. By "unintended consequences," I mean the effects of policy reform not included in the original stated aims. Some studies frame unintended consequences as "vulnerability to surprise," and others frame the consequences in "evaluative terms, arguing for claims that particular decisions led to outcomes variously described as unintended" (6, 2012). I will present studies that demonstrate situations where teachers face unintended consequences of reform.

One qualitative analytic project by Anderson, Butler, Palmiter, & Arcaira (2016) examines teachers in eight case study districts that faced early implementation of teacher evaluation reform. The teacher evaluation reform intended to improve student test scores ultimately, which most of the teachers in the study recognized. While teachers and administrators



reported that feedback from administrators in the new teacher evaluation system improved their instruction, no administrator would offer the conclusion that student achievement was improved by this reform yet. Some unintended consequences of this reform were positive. For example, the teachers felt the process of receiving feedback made them stronger and more reflective. One teacher wrote, "I do think that thinking about these elements [of the observation rubric] has improved my teaching." Another teacher wrote that it "completely changed the way I taught, the way that I viewed my classroom, the way that I reflected, and the way I saw my students." Another teacher remarked that she was better able to stimulate high-order thinking skills through improved questioning gained as a result of the new teacher evaluation process. Other unanticipated consequences were a surprise. As a result of the new teacher evaluation system, some teachers requested a transfer to a less challenging school, thinking they would get a better evaluation score with what they perceived to be an easier population of students to work with.

Sometimes policy does not play out as expected. In Tennessee, the RTTT legislation requires that half of a teacher's evaluation come from observations and the other half come from student achievement data. "In a bizarre twist on the notion of accountability, teachers in untested subjects and grade levels must choose the scores of another teacher to count for their evaluation." (Darling-Hammond, 2012, p. 135). New York City Public Schools faced similar issues in the use of value-added measures which incorrectly assigned data to teachers and resulted in erroneous evaluations (United Federation of Teachers, 2010). In Tennessee, the amount of time an evaluator had to spend in the formal observation cycle with a teacher ended up equating to a whole day per teacher when considering all the elements to be completed: the pre-observation conference, the actual observation, the inputting of data, and the conducting a post-observation conference. So, unintentionally, the amount of time this new system took was huge. One

principal said of the new Tennessee evaluation policies that they“ put everyone under stress, are divisive and suck the joy out of building” (p. 135). This study highlights the need to pilot teacher evaluation reform in order to acknowledge issues and to study the plan’s drawbacks before fully implementing it across a whole district.

### **What Does Teacher Research Reveal About Teacher Evaluation Reform?**

Should teachers’ input be incorporated into the development and rollout of reforms? The concept of the etic, or outsider, perspective in anthropology may provide some insight regarding this question. The research shows that an etic analysis can contribute to a locally held knowledge system, but etic knowledge should stay in that setting and not be applied to other settings. As Heath and Street (2008) argue, ethical questions arise when etic knowledge is introduced to other settings. The value of this “sharing” is best left up to a researcher and subject in a collaborative effort. Research about teachers’ knowledge could be valuable in developing local policy, but it may not work in a different setting or when applied universally across the country. Some argue that the researcher has a responsibility to include all participants, including the teacher, in their research ideas. Cochran-Smith and Lytle (1993) argue that “teachers’ roles as theorizers, interpreters, and critics of their own practice” have been ignored and “they have rarely been included in decisions about research as knowledge generation” (p. 1). They argue that researchers’ work is still seen as the most valuable source of validity for teachers’ professional practice.

Cohen, Moffit, and Goldin (2007) argue that rather than policy dictate practice, policy should rely on it. They examine the reliance of policy on practice, pointing out that there have been numerous grant programs and incentives for education reform efforts, but the returns on investment have varied widely. Because offering programs and funds has not worked to create

educational change, Cohen, Moffit, and Goldin (2007) argue that policy actually depends on practice. They write, “policies will be more likely to influence practice if they cultivate practitioners’ use of policy and less likely to influence practice if they only allocate and regulate resources” (p. 517). The researchers argue that the practitioners are the key problem-solvers to which policymakers should turn to direct their efforts. Teacher research presents a promising methodological vehicle to convey teachers’ work to policymakers.

### **Teacher Research**

The field of practitioner inquiry stretches back over a century (Lagemann, 2000). Yet, only in recent decades with the work of Stenhouse in the 1970’s, Carr and Kemmis in the 1980’s, and Elliot and Noffke both in the late 1990’s among others, have researchers formalized it into a movement, claiming its place in the field of education as a respected framework for scientific research. While there are several different types of practitioner inquiry including action research, self-study, and the scholarship for teaching and learning, this research study focuses on the use of teacher research.

Teacher research was popularized in the United States in 1980s (Lagemann, 2000) and examines issues in teacher learning, school reform, and teacher’s knowledge. Teacher research can have implications that go beyond the school site; it works well with professional development and teacher preparation. However, teacher research should not be generalized to create global theories or strategies for all contexts in a way that neo-positivist research attempts to accomplish. The Brookline Teacher Research Group, Philadelphia Teachers Learning Cooperative, and Prospect Center teachers’ group are examples of established communities of teacher research.

Teacher research assumes the professional context is a site for inquiry with blurred boundaries between research and practice. Teacher research groups may be formed among

teachers or with the assistance of a university partner. Teacher research projects often focus on inquiries related directly to teachers' work, such as student learning. One typical teacher research study is described as follows: four early childhood educators formed a grant-funded research group to learn more about their students' literacy and language development across different care centers. Their work was accountable to the grant guidelines, but they came together due to their interest in teacher research as a form of professional development. They conducted monthly meetings, identified case study children, kept field notes, arranged cross-visits to each other's schools, read articles together, participated in conferences and finally disseminated their findings through literature. They remarked that the participation in teacher research helped them "move beyond the traditional physical and professional boundaries of [their] work" and combated the isolation they felt (Espiritu, Meier, Villazana-Price, & Wong, 2002, p. 78).

In teacher research, validity is established through "significance, quality, grounding, and the authority of autobiographical analysis" (Bullough & Pinnegar, 2001 as cited in Cochran-Smith & Donnell, 2006). Analysis of data is primarily orally constructed and storytelling is a major feature of the research, as these methods are akin to teachers' work.

I will present three cases where teacher research has influenced policy. Rust and Meyers (2006) researched fellows in their Teachers Network Leadership Institute (TLNI) which engaged in teacher research that resulted in policy changes at their schools and districts in the 2000s. The following examples illustrate this point: Fairfax County teacher research resulted in a new teacher leader certification, teacher research in Los Angeles influenced their union's salary point guide, and New York Teaching Fellows teacher research has been integrated into funding decisions. As a result of the research, teachers reported that they improved their understanding and skills related to policy and some teachers in Wisconsin became political advocates for their

findings. Kono-Lungberg (2013) explains how the New York Metlife Fellows worked together to influence student achievement and economics in New York City public schools. They created a publication called *Making the Case!* to amplify teachers' voices in the policy arena. These teachers' professional development through their teacher research was valuable in its own right, but it amplified through the TLNI network in a way that connected it to and provided an influence on policy.

J. Myron Atkins (1988) describes a teacher research project that influenced policy and had a meaningful impact on its participants in California during the 1980's. Seven mentor teachers from four different school districts participated in a grant-funded teacher research study with professor and researcher J. Myron Atkins as the facilitator. The mentors were concerned about threats to the program in which they participated, gather research for their concerns and produced a publication called "What Are Our Worries, Hopes, and Dreams for the Mentor Program?" They circulated 4,000 copies of this monograph to the Department of Education, members of the legislature, and neighboring school districts. It is due to this research and publication that the teachers believe their mentoring program was saved.

Fregeau and Leier (2002) discuss the methods they used to influence policy in a teacher research model called Teacher Envisioned Research and Reform Plan Approach (TERRA). TERRA's mission is to help educators realize that they have power to initiate socially-just reform with an ultimate goal to reduce inequities among children. Participants used dialogue journals, discussion, and cross-cultural activities to understand their social reality. Then, they conducted a meta-analysis of those activities. The researchers used examples of teachers creating change in educational policy to help participants realize their social reality. Then, researchers guided participants to imagine how to create a reform plan, gather ideas for reform and take

action to implement their ideas. TERRA showed that participants experienced a variety of responses including “retreat, anger, frustration and also a new determination” when they tried to overcome barriers to reform and implement their ideas (Fregeau and Leier, 2002, p. 181).

TERRA found that working with a university partner, who approached administration with ideas alongside them, were more successful than those who did not.

### **Conclusion**

The studies above represent a selection of research that focus on understanding current teacher evaluation policy mandates through policy analysis, exploring the effects of reform on teachers, and the use of teacher research in this project. Various methods of teacher evaluation have been researched extensively, including the use of student growth measures, peer assistance and review (PAR), observation, and student evaluation of teachers. The use of formal observation is widely used across all states and about half of U.S. states use student test scores in evaluating teachers (NCTQ, 2019).

What does the research reveal about best practices in teacher evaluation? The National Education Association (NEA) suggests a peer model of evaluation embedded in teacher-led schools (NEA, 2012). Based on findings from their quantitative study of Virginia’s teacher evaluation system, Xu et al. (2016) suggests that process standards such as planning, assessment, and professionalism are more predictive of student achievement program than formal observations of teacher practice. Peer Assistance and Review (PAR) is a 30 year old method that uses a collaborative process for evaluation among the teachers’ union, district administration, the board of education, and consulting teachers to provide support and evaluate teachers; however, it has not been widely studied (Harvard Graduate School of Education, n.d.). Studies about observation show that it should be conducted four to five times per year to ensure reliability and be accompanied by artifacts (Milanowski, 2011). Danielson, Marshall, Mid-Continent Research

for Education, and McREL Learning, Marzano, and Stronge all offer rubrics for use during observations. The use of student evaluations to rate teachers is understudied, but results show problems with this approach (Polikoff, 2015).

Current policy approaches show that depending on the view of teaching, administrators take on different roles (Darling-Hammond, 1984) in teacher evaluation. Additionally, multiple measures are promoted and teachers are put on a continuum from initial certification to master teacher (Darling-Hammond, 2012; NEA, 2012). In New Jersey, the TEACHNJ Act was introduced as “historic tenure reform” according to its official text and introduced a reliance on student outcomes in adherence to federal mandates set forth by the Obama administration’s Race to the Top waivers and eventually the reauthorization of the Elementary and Secondary Education Act in 2016. Local New Jersey districts had to comply with this state law, and they updated their policies around the same time. Teachers in New Jersey are evaluated by observation and student growth goals.

The research shows that reform can have negative effects on teachers. This includes effects on teachers’ sense of self, low morale, anxiety, and dread (Evans, 2000; Troman, 2000). Yet, there are also certain factors that help teachers build the capacity for change. This includes teachers’ high level of metacognition (van den Berg and Schulze, 2014), feeling satisfied with their jobs (Evans, 2000), and resisting nostalgia (Goodson et al., 2008). Low trust of teachers (Troman, 2000) and lack of comprehensive supports (Ingersoll, 2001) can lead to “burn out” and affect teacher retention (Robertson-Kraft & Zhang, 2016).

Unexpected and unintended consequences of reform often occur. Sometimes these are positive, such as teachers who reported improvement to their practice (Anderson et al., 2016) while much posed further challenges. Anderson et al. (2016) report that teachers requested

transfers to less challenging schools as a result of the teacher evaluation changes in their district. In other examples of the rollout of new policy, teachers in untested subjects had to choose other teachers' data to use (Darling-Hammond, 2012), data was erroneous (United Federation of Teachers, 2010), or the amount of time observation took equated to a whole school day per teacher (Darling-Hammond, 2012).

Teacher research shows potential in revealing and addressing the consequences of reform, such as teacher evaluation policy mandates. Research shows that policies that simply allocate resources are not effective; policies that engage and develop practitioners are better (Cohen et al., 2007). Three case studies of teacher research influencing policy include Kono-Lungber (2013), Atkins (1988), and Fregeau and Leier (2002). In this chapter, I focused my literature review on understanding current teacher policy mandates through policy analysis, exploring teachers' responses to policy, and the place of teacher research in this effort.



### III - METHODOLOGY

I approached this research using the conceptual framework of “teacher research” (Cochran-Smith and Lytle, 1993). Cochran-Smith and Lytle define “teacher research” as the “systematic, intentional inquiry by teachers” found within a community of discourse (p.5). The authors argue that the trends of research take an “outside-in” approach where knowledge moves from university or policymakers to teachers.

There is an underlying view in policy-dominated perspectives of teacher education today that teachers cannot fix themselves, but need the molding of an outsider, such as an educational consultant (Cochran-Smith & Fries, 2001). My research shows that teachers have their own specialized knowledge grounded in their daily work and are operating within an educational context that frames their experience. Additionally, my research emerges from my own activist practice where I have worked within the system to create change in the contemporary social and political climate. I am inspired by the work and spirit of Alice Miel, who was a teacher researcher during the 1960s. In her time, she saw teachers as “grassroots curriculum makers” who were replaced with “psychologists, subject matter specialists, and national efforts to supply America’s schools — regardless of their location or participation — with the best curriculum money could buy” (Schubert, Marshall, Sears, Allen, and Roberts, 2007, p.72). She fought back with a moral imperative to empower teachers who were local experts immersed in the context.

In my quest to value and amplify the voices of K-5 female elementary school teachers in policy, this qualitative study explored what a collaborative teacher research project can reveal about new teacher evaluation mandates. In order to understand and explain the effects of the new teacher evaluation reform on teaching in their school district and document the use of teacher

research to engage teachers as policy actors, I posed the following research questions in a collaborative teacher research project:

1. What are the underlying discourses in the official documents of teacher evaluation policy?
2. What responses surface when teachers study teacher evaluation policies and practices in their New Jersey school district?<sup>9</sup>

### **Overview of Study Design**

This qualitative research project comes out of the field of practitioner inquiry, specifically under the heading of “teacher research.” Cochran-Smith and Lytle’s (1993) teacher research assumes that teachers’ daily work provides an experiential base for their knowledge about education. The goals of teacher research can be powerful, as there are “different emphases on individual and institutional growth and on the promotion of teacher research as a means to problem solving, to technical improvement, or to strategic social change” (Cochran-Smith and Lytle, 1993, p. xiv). A teacher researcher’s agenda depends on her position in the school and can “further different agendas or outcomes” (Cochran-Smith and Lytle, 1993). Researchers may automatically think that teacher research is only about inquiries related to the classroom, but it has been conceptualized to go much further than that. The broader agenda of Cochran-Smith and Lytle’s (1993) “teacher research” moves beyond the classroom and focuses on “educators’ sense of social responsibility and social action in the service of a democratic society” (p.58).

My methodology choices are heavily influenced by the work of Sandra Hollingsworth (1994), who outlines a feminist approach to teacher research, including conversation, self-

---

<sup>9</sup> This question originally referred to the phenomenon as “teacher evaluation reform” but, for reasons that became clearer during the course of the study, changes in policies and practices were actually central to the investigation rather than changes signaled by reform. I discuss reforms in Chapter 3 more fully.

evaluation, critique, and a commitment to social change. In a conversational piece with fellow researcher Janet Miller, Hollingsworth teaches us to ask ourselves, “What is the valuable knowledge these participants have that I don’t have?” (p.130). In her research, she found that teachers experience a loss of self-esteem and are always trying to measure up to someone else’s standards without success. The teachers she researched never felt good enough. Her findings came out of narrative methods like the ones suggested above. I solicited my participants’ responses to reform using narrative methods, as well.

### **Researcher Positionality**

I came to this research as both an insider and an outsider. I consider myself in a liminal position where I am both a full-time elementary school teacher and also not just a teacher at the same time. On the one hand, as a fellow teacher at the research site, I occupy parallel positions as my participants in the school hierarchy by sharing the same principal, educating the same community members’ children, and facing the same day-to-day teaching responsibilities. This also put me in a position to be heavily involved in the discourse around the reforms I am exploring through my colleagues’ eyes, having my own thoughts and reactions to the research questions that I acknowledged and tempered as I collected data. We all have different relationships with the administrators, and mine tend to be friendly and collegial.

I was also an outsider in this research project. I have gained some distance from my perspective as a teacher at the site through my leaves of absence following the births of my daughters and by engaging in other work, such as supervising, working in politics and the teachers’ union, and by studying as a doctoral student. My relationship to some of the teachers in the study was somewhat new as I was just transferred to the school in 2018 after having taught in the district for many years at another school.

Additionally, I acknowledge my position of authority in this research project. I designed the study, chose the topic, and called teachers together to investigate my research questions. This was challenging for me as someone who has taught for nearly two decades, and I truly wanted to see myself as a participant on an equal footing with the other participants. However, using that term “participant” masks the power and agency I had while conducting this research. Thus, although I wanted to engage in a teacher research group as an equal participant, it is evident that without my initiating the research project, the teacher inquiry group would not have existed.

In qualitative research, “eliminating the actual influence of the researcher is impossible and the goal in qualitative study is not to eliminate this influence, but to understand it and to use it productively” (Hammersley & Atkinson, 1995 as cited in Maxwell, 2012). The research participants are always influenced by the researcher, whether in something as obvious as the use of leading questions or simply the physical presence of the researcher. So, it is crucial that I, as the researcher, stay aware of the context of knowledge construction and how I am influencing the participants in discussions. Again, the goal is not to eliminate influence, as it might be in an experiment, but to recognize and reflect upon it. For that reason, my researcher’s journal was crucial to provide reflexivity while working with the participants in this collaborative teacher research project. In the journal, I continued to capture my own thoughts at each step of the research process, especially to reflect upon on my actions as a lead teacher-researcher and to evaluate my job as a facilitator of a what I intended to be a collaborative teacher research project. The journal served as a place for what Schön (1983) calls “reflection on action” after the event has occurred.

## **Pilot Study**

In May of 2018, I conducted a pilot study where I led a group interview. I included myself and two K-5 female elementary school teachers regarding their responses to a recent mandate. The participants had been teaching 10 or more years across various school districts and chose to focus on teacher evaluation first, as it was meaningful to them. However, I instructed them to pick a different topic since my study would be on teacher evaluation, and they settled for their second choice of standardized curriculum as their topic of inquiry. The interview took place for 1.5 hours at a location outside of the school district. I wanted to explore how these teachers made meaning of policy in general, as well as noting their reactions to the specific policy and procedures as related to standardized curriculum.

The pilot helped me formulate my study design, particularly for planning and facilitating the teacher inquiry sessions. I noticed that participants were open and shared a great deal starting 15-20 minutes into the session, speaking both positively and negatively about the topic. For that reason, I included warm-up type activities in each of the teacher inquiry sessions. I found that the teachers had limited-to-no understanding of what a policy is, how the term is defined, and how to use the term in a sentence. They were unfamiliar with words I used in my questioning, such as “landscape” and “teaching context.” Based on that finding, I decided to provide an overview of policy in the first session of the teacher inquiry group. Following the suggestion of Ashendon, Connell, Dowsett and Kessler (1987), “teachers’ interest and understanding of educational reform is an important strategic issue and needs to be encouraged and developed,” (as cited in Fregeau and Leier, 2002). Previously, I assumed that my colleagues would know as much as I did about policy. Lastly, I found that I gained richer data from a drawing exercise than just a discussion about the topic. So, I decided to include similar exercises in the teacher inquiry group.

In addition to these findings, I also grew as a researcher. I developed a more conversational approach at facilitating discussion because of my reflections on the pilot. I also discovered that the participants' narratives were interwoven in the discussion rather than set aside to illustrate a point. I looked for these narratives appearing in this format when I conducted data analysis.

## **Research Site and Selection of Participants**

### ***Research Site***

The research site was chosen primarily due to my relationship with the community and participants. As a practicing elementary school teacher, I knew that these relationships would give me the most access and deepest answers and would provide rich data for my research. I was also cognizant that I needed to monitor my own bias as I conducted the project and took measures to do so, which I detail later in the proposal.

The general attributes of the school district that was my research site are as follows: a mid-size, public elementary school situated in a mixed rural-suburban community in New Jersey about one hour from New York City. Many people characterize New Jersey by its proximity to nearby New York City and Philadelphia, focusing on the human made, industrial, and commercial landscapes found in its population centers, such as Newark, Atlantic City, and Camden. Yet, the Garden State is diverse, with a variety of geographical features, including coastlines, forests, farms, and mountains. The school district's community is located in the Appalachian Ridge and Valley Region, which is rich with farmland, open space, and settler colonial history. Most of the students reside in houses that have been erected on developed farmland, many with large acreage surrounding them, and some students live in old colonial homes near the historic centers of the town. Farms still dot the countryside, selling trees, fruit,

vegetables, eggs, and meat. Many of the students participate in extracurricular activities, especially sports, youth groups, dance, theater, and religious activities.

The student population of approximately 1,500 students is predominantly white and English speaking. Eighty-five percent of students are white, 8% are Hispanic, 4% are Asian, and 1% of students are Black. In terms of languages, 96% of students speak English, 2% speak Spanish, and 2% speak Slovak, French, Urdu, or other languages. Seven percent of students receive free or reduced lunch. The median home cost in 2022 is \$452,300. Few to no students live in affordable housing as there is a lack of availability in the region; however, with the township's expansion of housing developments with affordable options, 95 or more students who may or may not living in affordable housing are expected to join the district in the next five years. The district generally meets or exceeds the expectations of learning as compared to the statewide average performance on state assessments.

According to the National Center for Educational Statistics (NCES), in 2017-2018, 70% of teachers in New Jersey were reported as under 50 years old, and 81% of teachers identified as female. This school district reflects the typical public elementary school in New Jersey in terms of gender, with the majority of teachers identifying as female and under 50 years old — at 94% and 74% respectively. Additionally, the majority of teachers in the school are white, as shown in the NCES's most recent 2017 national data on mid-size elementary schools. The wealth of the teachers does not match the median household income in the municipality. Teachers in the district earn between \$40,000 and \$80,000 per year, while, according to the latest 2021 Census data, the median household income in the municipality is around \$122,000. Few teachers live in the district and most live in neighboring towns that are less expensive, or in nearby Pennsylvania, where the cost of living is more affordable.

In the context of teacher evaluation reform, the district adopted a new teacher evaluation system in 2014 with revisions in 2016 and 2017. This timing matched New Jersey's adoption of the Teacher Effectiveness and Accountability for the Children of New Jersey Act (TEACHNJ) and AchieveNJ, keeping the district in compliance of the state laws New Jersey Administrative Code N.J.A.C. 6:10 and New Jersey Statutes Title 18A (N.J.S.A. 18:A). N.J.A.C. Educator Effectiveness, commonly known as "Chapter 10," outlines the rules for school districts regarding minimum requirements of teacher evaluation in New Jersey. N.J.S.A 18:A, known as TEACHNJ, describes the law around teacher tenure and dismissal. As stated earlier, the district has been utilizing rubrics based on Charlotte Danielson's *The Framework for Teaching* (2007, 2011, 2013).

### ***Participant Selection***

The participants for the research study were my colleagues in the same school district. I recruited a purposeful sample (Maxwell, 2012) of at least three experienced in-service, K-5 teachers in the same school district who identified as female. I was able to recruit a beginning teacher and master teacher as hoped (see Appendix C and D). These participant categories are explained further later in this section. Koro-Ljunberg (2014) offers a framework for conducting policy-oriented teacher inquiry and suggests that sample sizes fewer than 25 "can create a sense of community and activate participants as a group or collective" (p. 774). Since this is a collaborative teacher research project, I promoted that community to mine richer data regarding the study's research questions.

In my study, I treated gender not as just a background fact but as a condition that cannot be divorced from the meanings of my findings. I chose to focus on K-5 female teachers for several reasons. First, the typical population of elementary school teachers in the United States



identify as female and white. According to the most recent NCE 2008 national data on mid-size elementary schools, 78% of public elementary teachers are female, and 85% of teachers are white (U.S. Department of Education, 2012). Second, there is a call in the world of academia to explore this population's experiences, since their perspective is often trivialized and understudied (Galman and Mallozzi, 2012). U.S. elementary school teaching is a "site of more intensified and institutionally complex professional feminization than any other level of schooling," and it is often thought of as "domestic labor" (Galman and Mallozzi, 2012, p.247). According to a literature review by Galman and Mallozzi (2012), there is a need for more research on these women's experiences. Teaching is gendered work and taking steps toward professionalization makes steps toward equality and control over women's own work in the face of rationalization and standardization in reform efforts (Zeichner, 1991). Finally, I chose to focus on teachers who are actively teaching rather than pre-service or retired teachers. While pre-service teachers may visit to observe or conduct their student teaching internship, and retired teachers may visit as guests or act as substitutes, active teaching faculty who are employed as such in the district are the most central actors in this research setting. While it would have been interesting to learn from non-tenured teachers' perspectives, I believe the risks posed to non-tenured teachers was too great for them to speak freely in this study, so I only recruited tenured teachers. For these reasons, this purposeful population comprised K-5 school teachers who identified as female and who were actively teaching and performing the other duties required of staff.

I recruited at least three tenured in-service, K-5 teachers in the school district who identified as female and who were actively teaching and performing the duties required of staff. As a female K-5 teacher who has worked in the school district for 18 years, I participated as an additional teacher in the experienced participant category of this collaborative teacher research

study. I recognize that my position as researcher limited my ability to participate on equal footing with the other teachers. I recruited participants by verbally inviting teachers who had registered for my teacher inquiry group. I asked these teachers in person or over the phone if they would like to participate (see Appendix D).

The teachers who participated in this study had tenure, or job security, and belonged to a teachers' union. Tenure is a safeguard to employment protected by our teaching contract and defended by the teachers' association. Recruiting only tenured teachers eliminated any unlikely risks to employment status.

All of the teachers in this study were white, middle-aged cisgender women who worked in the same PreK-8 school district at the time of data collection. Eleven teachers whom I have named with pseudonyms, Kimberly, Brittany, Dorothy, Melissa, Lauren, Nicole, Naomi, Teresa, Rose, Mabel and Tracy, participated in the teacher inquiry group sessions. Five of those teachers, Kimberly, Brittany, Dorothy, Naomi and Teresa, also participated in follow-up interviews.

At the time of data collection, all of the participants had been teaching elementary school at least ten years, except one teacher who had taught for four years at the time of data collection. There was also one late-career teacher who started teaching forty years prior to the time of the study. All of the teachers had experience teaching in more than one grade level or assignment in their career and with more than one principal, supervisor, and superintendent. Several teachers taught in other schools before getting a job at the current district, including public and private schools. The teachers in the study were elementary teachers working in different schools within the same elementary school district who joined the study with the incentive of earning professional development hours toward paid time off along with other teachers who attended the session but did not participate in the study.

Participants were informed that given that this is a study specifically for K-5 tenured, female teachers working in one school district, material gathered from the teacher inquiry group, observations, and interviews would be attributable to a relatively small group of individuals. Thus, it is possible that members of the school and district community would be able to infer participants' identities and/or surmise the source(s) of some opinions contained in presentations, reports or publications coming out of this work. Participants were reminded of this risk again during the study's activities if they shared or recommended something controversial.

As the lead researcher, I took measures to minimize these risks by: (1) recruiting a larger pool of participants in the teacher inquiry group than those actually participating in the study, projected to be twice the number of participants based on past enrollment in professional development workshops; (2) not revealing participants' identities to anyone inside or outside of the school district and instructing participants to do the same; and (3) relying on paraphrase rather than direct quotations in my notes when possible.

Additionally, teachers' participation in the discussions about teacher evaluation could have caused them to feel vulnerable when sharing personal stories with each other—a feeling many people experience when they talk about themselves.

The following table illustrates how the study was designed to recruit at least one beginning teacher, defined as having tenure less than seven years, and two or more experienced teachers, with at least one master teacher who has taught more than twenty years. I based these participant categories on the introduction of new teacher evaluation systems. For example, a beginning teacher would have received tenure in the newest teacher evaluation system, while the experienced teachers would have received tenure in previous teacher evaluation systems. Earning tenure under a different system offers a different perspective to what teachers face today.

**Table 2***Participant Categories*

	Years Tenured	Target participants	Actual participants
Beginning teachers	<7	≥1	1
Experienced teachers Master teachers	8 < x < 20 >20	≥2 ≥1	10 1

**Data Collection Methods and Analysis**

While the attempt to professionalize teachers' work has been a big push of teacher research, researchers must avoid what Zeichner (1991) calls "the intensification of teacher's work," so time must be built into the school day to participate in teacher research. Therefore, my goal was to keep the research in the existing professional development structure as to not add additional responsibilities to teachers that may take time away from their instructional duties.

I discussed this project with administrators verbally and then followed the current protocol, which is to request approval from the current superintendent. The district is a friendly environment for teachers' continuing education and professional development. Yet, in anticipation of any concerns around the project taking time away from teaching duties, I proactively addressed this concern by utilizing and supporting an existing professional development structure outside of the regular school day. It is important to the administration that any district-supported professional development activities fit within the district's agenda. Therefore, I pointed out the connection that our storytelling activities can be adapted by teachers to enhance the content-area instructional delivery in their classrooms.

Data collection lasted from October to January. There were multiple sources of data collection to provide triangulation and increase the validity of the study, including a (1) teacher inquiry group focused on storytelling and oral inquiry; (2) a document analysis of official policies; (2) informal observations; and (3) follow-up interviews. The connection between the data collection methods and my research questions is shown in the table below, with a more in-depth explanation of each method following.

**Table 3**

*Data Source - Research Question Correlation*

Research question	Data collection plan	Rationale
1. What are the underlying discourses in the official documents of teacher evaluation policy?	Researcher-led document analysis of teacher evaluation policies	Explore underlying discourses in official documents.
2. What responses surface when teachers study teacher evaluation policy in their school district and state?	Researcher-led discussion of teacher evaluation policies and teacher participation in storytelling activities during the teacher inquiry group. Some teachers also engaged in follow-up interviews with the lead teacher researcher. These were audio recorded with the participants' permission, transcribed, and analyzed.	Explore teachers' meaning making of the policy, their feelings and attitudes, and the framing of their professional identities in light of teacher evaluation.

### ***Teacher Inquiry Group***

The teacher inquiry group sessions' protocol was heavily influenced by the creative ideas and specific, arts-based methods utilized by Lee Bell's (2010) Storytelling Project. Storytelling, discussion of hopes and fears, and dramatic representation are examples of specific methods in the protocol to help access teachers' responses to teacher evaluation policy. Word clouds (Stapleton, 2018) were also used to provoke discussion. Stenhouse (1978) envisioned teacher research as teachers writing case studies of their own classroom to influence policy (Somekh & Zeichner, 2009). I gave the teachers what they needed to participate in the discussion about policy and collect data about their own experiences. See Appendix B for Teacher Inquiry Group Protocol.

The teacher inquiry group utilized an existing professional development structure in the school district. Teachers sign up for workshops, which can range from book clubs to curricular presentations, in exchange for "flex time" during the school day. Use of this existing structure illustrated my willingness to work within systems to create change. It also created a mutually beneficial relationship where the district can benefit from the research rather than creating additional tasks or what may be seen as distractions for hardworking teachers.

This group met three times over the course of the Fall-Winter for approximately 1-1.5 hours each time. I was a full member of the group, but also acted as a lead researcher who scheduled the meetings, took notes, and facilitated discussion. Teacher research has been found to be more successful when there is a university partner (Fregeau & Leier, 2002), so I acted as such during the inquiry sessions in addition to my role as a colleague.

Digital recordings of teacher inquiry sessions were made, reviewed, and protected in a local storage service only available to me. Pseudonyms were used in any publications arising out of this work. The teachers' professional backgrounds remain unexplored, and their comments

have been combined under themes to help protect their confidentiality. Participants consented to these terms as part of the consent form (see Appendix C).

### **Inquiry Session 1**

The first meeting of the teacher inquiry group made space for participants to share recollections of their formal observations, define what made them feel pleased and successful as teachers, share their hopes and dreams for teacher evaluation, and offer up any recommendations for policy.

This session was designed with the intent that teachers would share their recollections of teaching before, during, and after the introduction of the new policy. While the teachers spent time sharing their recollections of the first time, they were each formally observed for evaluation purposes, all recollections except one predated the introduction of the new mandate.

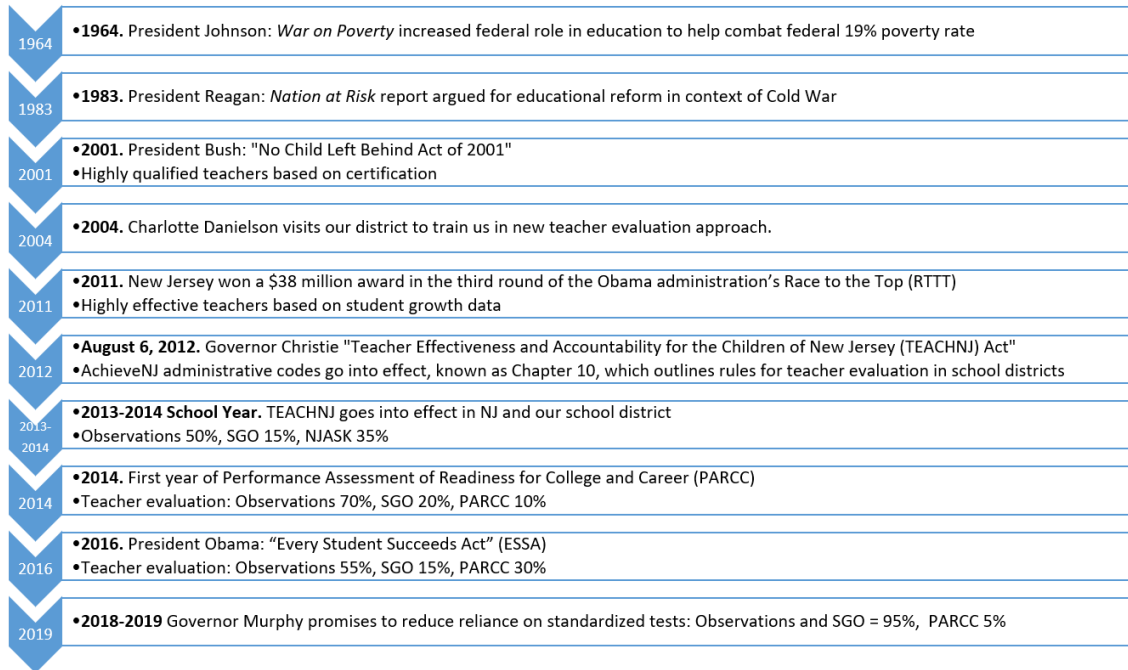
Teachers also co-described the logic, purpose, and structure of the new teacher evaluation policy. Findings from the pilot study confirmed the arguments made by Ashendon, Connell, Dowsett and Kessler (1987) that teachers' interest and understanding of educational reform needs to be encouraged and developed (Fregeau and Leier, 2000). For this reason, the main focus of the first meeting assisted teachers in entering the unfamiliar terrain of policy, equipping them with the domain-specific terms they needed to participate in the educational reform discourse.

I provided a timeline teacher evaluation policy in New Jersey in the format of a twenty-minute long PowerPoint presentation. This included a timeline of legislative and regulatory moves by the federal and state governments as well as by the school district (see Figure 4).

## Figure 4

### *Teacher Evaluation Policy Over the Years*

#### Teacher Evaluation Policy Over the Years



I presented the information as lecture due to time constraints but want to problematize the inherent "banking concept of education" (Freire, 1972) found in a lecture format. I facilitated a critical conversation among participants and stepped back so teachers could take up the space to analyze and challenge that which was presented. The teachers shared their reactions to this overview and also shared what makes them feel pleased and successful as a teacher. They engaged in group meaning-making through questioning and discussion. At the end of the session, teachers shared their hopes and dreams about teacher evaluation.

The first meeting focused on storytelling around evaluation, a lecture and discussion to equip teachers with pre-requisite terminology and concepts to further participate in the teacher evaluation discourse, and an exercise in dreaming together.



## **Inquiry Session 2**

The second inquiry session began with teachers rating how well certain statements resonated with them through what is colloquially called a “consensogram” among teachers, or a playful form of a Likert scale. Teachers placed sticker dots to express their agreement with each assertion in a continuum with “true for me” and “not true for me” on the extremes of the line. I chose a scale that valued the participants’ knowledge and experiences rather than upholding the idea that truth is objective. After all of the ratings were complete, I physically held up their compiled ratings, read the statement and used open-ended verbal prompts like “any thoughts about that?” to make space for further discussion. I chose the assertions at the top of each page based on findings from research about teachers’ reactions to policy and developed some to relate to the research questions of this study. The statements emerged from existing literature, as presented in Chapter 3, around teachers’ responses to reforms and were stated as follows:

- As a teacher, I have a say in policy.
- Current teacher evaluation policies are helpful for me as a teacher.
- Teacher evaluation accurately evaluates teachers.
- I am a critical thinker.
- I embrace change.
- Policymakers must trust teachers’ ability to teach.
- Teachers trust each other’s ability to teach.
- Supervisors trust teachers’ ability to teach.
- Supervisors trust teachers’ ability to teach.
- I feel anxiety due to my job.
- I feel satisfied with my job.
- I embrace change fully.
- Teaching in a school is like being a part of a family.
- I feel supported in my job.
- I am my biggest critic of my own teaching.

After the discussion, I encouraged teachers to share stories about a time that administrators taught them something through the teacher evaluation process. Then, we viewed

an image of one of Cezanne's painted trees with the notion that it is a particular tree but also represents a universal idea of trees. I explained, "You don't have to know this tree, but you know that it is a tree and you have probably seen a tree like this somewhere." Using that analogy, I asked if there were any universal themes or larger ideas that have emerged, in their minds, about teacher evaluation policy during our time together. Teachers shared their responses, and then I asked them to draw a picture of the teacher evaluation process and shout out some words as they worked to continue building our construction of knowledge about teacher evaluation. The use of Cezanne's tree was meant to stimulate their thinking about the topic in a new way and shift the mode of communication from just textual to include multiple means of expression, including graphical.

### **Inquiry Session 3**

We started the final inquiry session in a similar fashion to how we ended Session 2. Teachers shouted out words to co-construct a word cloud to capture the latest big ideas about teacher evaluation. This was intended to serve as an icebreaker and help us get back into the work from Session 2 quickly.

Nisbet et al. (1997) highlighted that end products of policy-oriented research produces recommendations for actions, and teachers can utilize their stories to interrupt or transform policy. I shared three examples from my literature review that illustrated instances where teachers came together, conducted teacher research and influenced policy.

Koro-Ljunberg (2013) recommends the production of articles, memos, opinion pieces, or blogs as action. I had intended to organize with the teachers to co-authored a design proposal for an action plan. However, the group decided not to act during the inquiry session. As a result, I

facilitated a discussion about teacher activism with follow-up interview participants (see Appendix E Project Timeline).

If action had become an outcome of the project, this teacher research project would have shifted to an action research project. However, that was not the case. As Hall (1981) explains, the methods in a teacher research project are experimental and should follow the local knowledge of the participants. Respecting participants is key (McTaggart, 1997), including their perspectives, expertise, and time (Stapleton, 2018). Action projects produce knowledge but also empower change and transform the participants (McTaggart, 1997; Fals Borda, 2001). I went into the research open to the possibility that this could either turn into an action project or remain teacher research.

Teacher inquiry group sessions were the heart of the research project, offering a site for co-constructing knowledge regarding policy documents, sharing lived experiences around teacher evaluation, and my hopes of planning/implementing an action plan, which did not materialize.

### ***Analysis of Official Documents***

While documents make things visible, the technical documentation itself is nothing without its social use (Grant, 2019; Prior, 2003). “The prime issues for the social scientific researcher is therefore to investigate how documentation functions in situated contexts, not least because such documentation is invariably used in ways undreamed of by its creators” (Prior, 2003, p.87). The aim of the document analysis was to investigate local policies within the context of greater state and federal mandates. It was a starting point for the other qualitative data collection in the study that explored the social use and function of teacher evaluation policy in an educational setting.

In the sections below, I describe the methods I used to conduct a document analysis of local school district policy documents in conjunction with other policy components in the teacher evaluation system. This was a step to define the teacher evaluation system as represented by official documents. In subsequent data collection and analysis, I moved into the work of presenting and analyzing the data from my study: teachers' lived experiences with teacher evaluation policy, including how policy affects their professional and personal lives.

### **Methods**

The methods utilized for this document analysis are those recommended by Dr. Patrick Love of New York University (2003). First, purposive sampling was used to compile primary policy documents associated with teacher evaluation in the district, state, and federal levels. I started at the local level and searched the district's database of policies for those that addressed the evaluation of teachers. I searched for key terms: "teacher evaluation" and "evaluation." Those local documents referenced the state law Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) and state statutes that govern teacher evaluation, particularly those governed by Chapter 10 law, New Jersey Administrative Code (N.J.A.C) 6A:10-1.1, N.J.A.C. 6A10-2.1 through 2.5, N.J.A.C. 6A:10-5.1 through 5.4, N.J.A.C. 6A:10-7.1 and 7.3. I followed the references to the New Jersey State Government website and searched the primary texts of these laws, codes, and regulations. While reading those, I became aware of research done at the state level to gather public feedback through taskforces and formal public meetings between 2011 and 2013, before the law was passed. I added meeting notes and presentations from the state's exploratory work into the pool of documents I was analyzing. Just as local policies referenced state laws and regulations, the state documents and meetings referenced the need to satisfy requirements in the federal Elementary and Secondary Education Act (ESEA) for funding purposes. At this point, I had followed the logic up from local policy to

the federal level governmental documents that summarized the legislation around ESEA and its reauthorization in 2001 and 2015. I also had followed the Obama Administration's Race to the Top (RTTT), which incentivized his policy priorities to states, while the legislature struggled to find common ground around the reauthorization of the ESEA. While local district policies are the focus of this chapter, their heavy reference to and reliance on state statute makes it incomplete to consider them in isolation. In fact, it became apparent that most of the decisions around how teachers are evaluated come from the state and federal government. After collecting and reading these interrelated documents, I catalogued them in a table for organizational purposes. In the table, I noted the source, date, author, significance, rationale, context, and meaningful quotations from the documents.

Next, I categorized documents further based on methods suggested by Lincoln and Guba (1985): solicited or unsolicited, comprehensive or limited, edited or unedited, anonymous or signed, spontaneous or intentional, and intended audience. After that, I identified patterns in documents. The documents were coded by what Geertz (1973) described as "a process of guessing at meanings, assessing the guesses, and drawing explanatory conclusions from the better guesses" (p.20 as cited in Stage and Manning, 2003). Finally, I conducted content analysis to understand underlying themes, assumptions, and beliefs.

I conducted an analysis of the documents first by looking at patterns and second by drawing themes from those patterns. I used focused coding to identify repetitive ideas that occurred more than twice. The occurrences could be connected by similarity, difference, frequency, sequence, correspondence, or causation (Hatch, 2000). Findings from the document analysis are presented in Chapter 4.

### **A Note about Documents**

Prior (2003) examines the anthropological use of documents through examples from ethnomethodologist Dr. Harold Garfinkel's work. Prior stresses the influence of organizational documents in social relationships and not just the content found in the document. For example, Garfinkel (1967) and his colleagues researched clinic folders and found that they were assembled with their use in mind, wording things in a way "such that it could always be shown that the 'right' things were done to the 'right' person at the 'right' time" (Prior, 2003). The content did not matter as much as the use, because "records originally designed for a set of purposes could be routinely used for quite different ends" (p. 51). The meaning of the documents depended on who was using them and the circumstances around their use. Additionally, access to the records and the right to make changes was restricted to certain roles in the clinical setting. She recommends "Never look at documents in organizational settings as isolated tools, but seek to discover how a document is linked into the wider information storage and retrieval system of which it will form a part" (p.87).

What can be learned from Prior's (2003) analysis of Garfinkel's work in clinical documents? Garfinkel's work provides a new approach to examining documents for social use and not just content. At the local district level, the teachers in this dissertation study were not aware of how to access the district policies on teacher evaluation and were also unsure of their significance when presented with policy excerpts. Their "text" was found in the conversations with administration, the templates they used to submit student growth objectives, the "write-up" of the formal observations, but not the policy itself.

As Prior (2003) describes in Garfinkel's work, "Social talk belonged in social work records, psychiatric talk belongs in medical records, and nurse talk belonged in nursing records" (p. 56). The same goes for different roles found within a school district. Teacher talk is not found

in these policy documents at the district, state, or federal level. These documents are not meant for teachers. So, what happens when a teacher, like me, examines them? The social use of these documents does not position teachers with power. The district policies are written so that policy documents reference state statutes and read like rules for the superintendent to follow without any locally decided rationale. The policies read like marching orders, or in other words, like documents that will be used to ensure compliance. A contrast exists between district policy documents those at the state level. There is also a level of compliance language within federal requirements in the state documents, but more research and rationale is provided. State documents give a sense that they are not only filling a federal obligation, but also demonstrating the unique landscape of the state, historical priorities, and taking ownership over the more details in the document. While district policies are used by officials (the superintendent, board of education, and their legal team) in school governance, the state level documents circulate in the hands of federal compliance offices, advocacy organizations in the state, state agency lawyers, politicians, policy experts, and, at times, the media.

### ***Informal Observations***

Teachers in the district meet regularly throughout the school year. As per negotiated contract, they are given a weekly common planning time to meet with colleagues while students are engaged in their special subjects. Additionally, teachers meet as a school faculty after school 1-2 times per month. I collected additional field notes during these two types of meetings for the duration of my data collection period. It was counterproductive to tape record these sessions, since that would influence the participants' willingness to share. Since I was an employee of the school, I also participated in the meetings.

Conversations occurred spontaneously between participants and me, as I work at the within the district alongside the teachers who were part of the study. I carried a researcher's

journal to maintain regular field notes. In situations in the field where I could not jot notes in a journal, I made quick audio notes on a recording device and recorded them in my journal later. My jottings were focused on teachers' demonstration of responses and/or lived experiences regarding any aspect of the district or state teacher evaluation system, including student growth measures and observations of their instruction.

I approached the informal observations with an open mind about the effect of teacher evaluation on teaching in our school. I collected field notes focused on recording when teachers demonstrated responses to any aspect of the district or state teacher evaluation system, including student growth measures and observations of their instruction. I used the constant comparative method to analyze the data and reflected on my researcher's bias in a researcher's journal.

Informal observations of these meetings provided a window into the effect of teacher evaluation on teaching and learning.

### ***Follow-up Interviews***

I conducted follow-up, semi-structured interviews with co-researchers from the teacher inquiry group (see Appendix A). Similar to the teacher inquiry sessions, the interview protocol has been adapted from Lee Bell's (2010) Storytelling Project. Participants were reminded at the beginning of the interview that they would be recorded and given the option to withdraw if desired. Digital recordings were made, reviewed by the researcher, and protected in a local storage service only available to me. Pseudonyms will be assigned in any publications arising out of this work. Participants consented to these terms as part of the consent form or did not participate in the interview.

It was anticipated that, in an interview, a teacher might bring up an idea or specific example of how she is affected by teacher evaluation in her work that is observable. This possibility included a classroom lesson, performance of a duty, or an artifact. I remained open to



the possibility that a scheduled observation of the teacher in action may follow from an interview, but that did not occur.

### ***Methods of Data Analysis***

Bogdan and Bilken (2016) define constant comparative method as “a research design for multi-data sources in which the formal analysis begins early in the study and is nearly completed by the end of the data collection” (p.262). The data analysis was iterative and ongoing to avoid a large pile-up of data at the end of the project. Findings in the early stages of data collection were analyzed before proceeding to the next stage. It involved coding texts, linking together themes, and generating analysis based on those themes. Memoing enabled reflection on texts, while the sparing use of exemplar quotes grounded and highlighted teachers’ voices in the work. Insights from that first analysis were utilized to inform the next round of data collection. This process continued until a strong understanding of the research questions emerged.

Regular memoing along the way helped me generate analysis of the data, methods, and my position as researcher. One of the goals of the teacher research project was to make sure the voices of K-5 elementary school teachers contributed to the research, so I utilized in vivo coding, or the use of actual words from my text, to name the themes, to keep the analysis grounded in the participants’ attitudes toward teacher evaluation (Corbin and Strauss, 2008 as cited in Bernard, 2011). I also used exemplar quotes sparingly that captured the themes, to stay true to the teachers’ voices.

I created a matrix for each research question to display analysis with supportive data sorted by teacher. I used Microsoft Word and Excel for coding and identification of themes.

### **Limitations**

A major affordance and limitation of this study was my familiarity with the participants. We knew each other in a work and, sometimes, social settings. This familiarity afforded me more

trust and potentially richer data. I thought it also might restrict how I conduct myself as a researcher. For example, I thought I may feel limited in what I can ask the participants. After listening to the first recording, I decided to change my facilitation style to better balance my collegiality with my role as a researcher. I also reflected on my bias regularly through a researcher's journal.

I present this sample of teachers as marginalized by teacher evaluation reforms. However, in another conception, these participants can be seen as holding dominance in society over other groups. White women benefit from privileges in society not afforded to persons of color and have historically worked in opposition to women of color. Given that the majority of the sample is white, it does not allow the researcher to explore intersectionality with issues faced by teachers of color.

While men hold a dominance in society, when it comes to the elementary teaching force, male teachers are statistically in the minority. Since there were no teachers who identified as male in the study, this did not allow the researcher to explore complexities, such as societal expectations and stigmas, faced by male elementary school teachers.

Gender is left unquestioned in this study as all the participants identified as cisgender female. But in reality, gender is a complex construct. There has been an increased awareness of the lesbian, bisexual, gay, transgender, and queer challenges to the male/female binary. I chose to describe the participants as "identifying as female" instead of simply "female" or "woman" to be mindful of our increasing cultural awareness, but the study does not unpack the construct of female. Instead, it actually hinges on the designation of "female" as it explores the feminization of the profession and associated marginalization.

Finally, this study does not include other members of the school community, such as parents, residents, students, and administrators. Empowering all the members of the school community to have a say would promote more democratic decision making than just teachers as long as certain groups are given the type of training they need to fully participate and not just reproduce existing hierarchies (Zeichner, 1991). Conducting a study about teachers and equipping them with the tools they need to enter policy discourse is a great first step but does not fully support a democratic society in the end since they are only one type of stakeholder.

### **Presentation of Findings**

My dissertation has six chapters. The first three chapters are revisions of the three chapters found within my proposal. The remaining chapters present an analysis of the data I have collected in relation to my research questions. Chapter IV tells the story of findings from official documents related to teacher evaluation. Findings from the teacher inquiry group sessions and interviews are found in Chapter V. Chapter VI communicates a discussion of themes and patterns across the set, and an explanation of findings as related to the research.

#### **IV- UNDERLYING DISCOURSES IN TEACHER EVALUATION POLICY**

In this study, I investigate ways in which teachers make meaning of teacher evaluation policy mandates. When I originally set out to conduct this research, I found that teachers in my pilot study did not have enough familiarity with policy to engage in a discussion about the topic, including how the term is defined and examples of educational policy at local, state or federal levels. Despite the policy mandates having a deep impact on their lives, teachers were not engaging with the policy texts.

In order to help teachers engage, I designed the first session of the teacher inquiry group to provide an overview of teacher evaluation policy and mandates since 2013. At the time, 2013 was just five years before the inquiry group met. I chose this starting point as a familiar shift for the teachers in the study, since this was the year all of the teacher evaluation procedures and rubrics in New Jersey districts changed to align to the newest law. However, the mandates around teacher evaluation reach much further back in time, as they were designed and implemented based on a theory of action adopted at the federal level with No Child Left Behind (NCLB) in 2001. Conducting a document analysis of official documents around teacher evaluation mandates helped me understand the mandates at a textual level and tell one possible narrative of how we got here.

##### **Federal Mandates: Theories of Action**

The federal focus on teacher accountability is based on political and economic fervor, including fear of falling behind in the global economy and fear of overspending on public education. Federal initiatives starting with Johnson's 1964 *War on Poverty* paint a patriotic dream that every child has the potential to contribute to the leadership of the United States. Public schools are described as "democratic social-sorting devices if all participants were given

equal chances to move ahead in the social system” (Spring, 1989, p.123). NCLB is founded on a continuation of a decades-long concern about the United States’ educational competitiveness with other countries, particularly performing behind other industrialized and non-industrialized countries on various standardized tests<sup>10</sup>. Legislators even threatened this would create a greater decline in the American way of life over time if the education system was not “fixed” (New Commission on the Skills of the American Workforce, 2009).

As the amount of federal money going toward education continues to increase each year, legislators demand visible results for the public expense (National Conference of State Legislators, 2016). Consequently, federal mandates target teachers, as public workers who are paid salaries, benefits and pensions from taxpayer dollars without consequence. Teachers are also targeted based on economist Hanushek’s (1992) frequently-referenced study that overlooks the impact of poverty on students to conclude teachers are the greatest in-school factor on student achievement. Policymakers hold an underlying belief that many teachers know how to teach well, but are too lazy or somehow unmotivated to do so. Policymakers include language of accountability in No Child Left Behind as a way to sidestep union-protections, like tenure, to fire teachers. They include punitive measures in the legislation that attempt to motivate teachers to act differently through methods like public shaming and threats of job loss. Obama administration’s Race to the Top (RTTT)’s waivers purported to be less punitive but, in the end, use the same approach. The authorization of the federal Every Student Succeeds Act (ESSA) by Congress also continues test-based accountability measures targeting teachers. After over a decade of these harmful mandates, national tests show no lasting improvement in student

---

<sup>10</sup> These tests include Programme for International Student Assessment (PISA), Programme for International Assessment of Adult Competencies (PIAAC), and the National Assessment of Educational Progress (NAEP).

achievement. Instead, the result of the accountability system over ten years is “low teacher morale, plummeting applications to schools of education, the need to recruit too many of our teachers from the lowest levels of high school graduates, a testing regime that has narrowed the curriculum for millions of students to a handful of subjects and a very low level of aspiration” (Tucker, 2014, p.2).

### **Underlying Discourses**

This section seeks to explain and interrogate some of the prominent discourses found in the official documents related to teacher evaluation mandates at the federal and state levels. It is important to note that district and state mandates sometimes seem independent, but they are actually documents of allegiance to the federal rules in order to be lawful and to guarantee receipt of federal funding.

Federal mandates operated under the premise that improving teacher quality closes the achievement gap, but, in reality, they act as a measure to punish teachers. They also purport to be documents that are crafted and receptive to educator input. However, in my document analysis, I found that these official documents are inaccessible to teachers, unresponsive to public comment, and contain standards that are federally, not locally, defined.

### ***Shifts in the Conceptions of Great Teaching***

The theories of action behind federal policies like NCLB target teachers to solve achievement gaps. One angle of attack on teachers is to challenge those who were placed in assignments that differed from their certification. The federal government narrowed the work of teachers through a new credentialing process at the university and state levels due to this federal mandate. The mandate also discredited thousands of teachers, forcing schools to send home letters to their families that their child’s teacher was unqualified. The New Jersey Department of Education’s website claimed, “A significant number of New Jersey students — more than one-

fifth—are taught by at least one teacher who does not hold certification in the appropriate area” (New Jersey Department of Education, 2015, p.6). Additionally, it stated that the whole equity gap as related to teachers with the highest out-of-field placement is driven by 10 districts, so “if the Department solves the equity gap in the top 10 districts, the gaps will be nearly eliminated” (New Jersey Department of Education, 2015, p.19). Two decades later, we see that all of the attention spent to try to publicly shame teachers and schools did not close the achievement gap, which persists today.

The federal government continued to target teachers, but the language used at federal and state levels to describe “great” teaching shifted from a focus on qualifications to effectiveness. In compliance with the federal mandates, the New Jersey Educator Equity Profile describes “highly qualified teachers” as those who are “(1) fully certified or licensed by the State, (2) holds at least a bachelor’s degree from a four-year institution, and (3) demonstrates competence in each core academic subject area in which the teacher teaches” (p. 4). The 2015 New Jersey Educator Equity Plan also defines excellent educators as “educators who are experienced, in-field, and qualified who are fully able to support students in getting and remaining on track to graduate from high school ready for college or careers” (New Jersey Department of Education, 2015, p. 13). It suggests excellent educators are “working towards high academic standards,” and are not first-year teachers or novices, (i.e., they have taught in their district for more than four years).

With the passage and implementation of the TEACHNJ Act in New Jersey and reauthorization of ESEA in 2016, these conceptions of teaching became codified with measures of student outcomes. NJ’s report “Excellent Educators for All” (2015) claims that pairing data about NCLB’s Highly Qualified Teacher placements with teacher effectiveness data increases

access to high quality instruction. In Appendix E.2 Performance Descriptors in New Jersey's ESSA plan of 2017, it identifies each of the four levels of teaching in great detail, stating that "highly effective" includes many specifications around organized and intentional planning, fostering a classroom culture with high expectations, clearly communicated instructional purpose, leadership in the community, exceptional performance on SGO, and well above average student growth scores on SGPs. In lock step with federal mandates, New Jersey's policy reports show evidence of a shifting conceptions of great teaching as just credentials to those that include teacher practice and student growth. Even so, teachers remain the target.

The text of TEACHNJ (2012) states the importance of linking teacher evaluation to student outcomes: "The New Jersey Supreme Court has found that a multitude of factors play a vital role in the quality of a child's education, including effectiveness in teaching methods and evaluations. Changing the current evaluation system to focus on improved student outcomes, including objective measures of student growth, is critical to improving teacher effectiveness, raising student achievement, and meeting the objectives of the federal 'No Child Left Behind Act of 2001.'"

TEACHNJ also states, "The goal of this legislation is to raise student achievement by improving instruction through the adoption of evaluations that provide specific feedback to educators, inform the provision of aligned professional development, and inform personnel decisions." As stated in the 2015-2016 Educator Evaluation Implementation Report, the state encourages the use of teacher evaluation data from AchieveNJ to drive coaching priorities and professional development to improve the quality of teachers. Data from this report claims that the retention rate of effective or highly effective teachers hovered around 95% in 2015-2017, the



first two years of AchieveNJ, signaling that it had a positive effect on the quality of the teaching in New Jersey. It also claimed AchieveNJ resulted in more helpful information regarding the evaluation process. Where before, teachers were rated as “satisfactory” or “unsatisfactory,” teachers now receive one of four levels of rating. It also offered outstanding teachers opportunities to coach others in their school to improve their teaching performance. It aimed to clarify expectations for good teaching, provided a common framework for conversations about teaching, and improve leaders’ ability to be instructional leaders. TEACHNJ also strived to support novice teachers, since there are still so many in the state (Excellent Educators for All, 2016).

### ***Punitive Logic***

When the department of education in New Jersey toured the state for stakeholder input in 2018, one participant’s feedback stuck with me during my research. They suggested, “Transition to an assessment that is less punitive for students and educators.”

Although one stated federal goal of the teacher evaluation system in Race to the Top (RTTT) is to provide a support mechanism for educators, in many ways, it operates as a system that monitors and punishes teachers instead. The punitive tone of this policy originated with No Child Left Behind (NCLB). Given that the bipartisan passage of NCLB did not occur until after the terrorist attacks of 9/11, the context in which this legislation passed was a nation in urgency and a presidential administration that operated in a tone of surveillance and punishment. The Obama Administration critiqued the “punitive labels” and “one-size-fits-all federal mandates” that were at the foundation of NCLB’s theory of action (ESSA, 2015), but followed the same logic in RTTT and the reauthorization of federal mandates around teacher evaluation.

In its Comprehensive Plan for ESSA, New Jersey noted in the subheading, “Capitalize on shift to ESSA by Focusing More on School Supports, Not Punishments” that “Stakeholders often shared the negative implications of the rigid and often punitive tone of NCLB” (New Jersey Department of Education, 2017, p.30). It was a stated intention of ESSA to be less punitive, but the approach persisted.

ESSA still required punitive thinking but with a softer approach in messaging. Minor terms like “annual yearly progress” were abandoned in favor of states’ designation for schools deemed not improving, in New Jersey’s case “schools in need of support and improvement.” Rather than having the federal government set goals, states were in charge of identifying failing schools and providing “supports” instead of punishments; however, if improvement was not observed, New Jersey noted that “schools and LEAs should be held accountable for the outcomes of their students” (NJ Consolidated Plan, 2017, p.33). Additionally, ESSA and the ESEA waivers that led up to it require states to use standardized tests to measure student outcomes and teacher effectiveness. These high-stakes measures have punitive consequences, which included state takeover of some operations, such as reassigning staff, reorganizing schools, and in some cases, closing schools.

There have also been indirect effects regarding the status and reputation of schools that do not have exemplary test scores, as they are downgraded on real estate websites and popular U.S. rankings. Even though ESSA touts that it encourages supports rather than punitive consequences, the punishments are still employed behind coded language.

Not surprisingly, since the state level mandates originate at the federal level, the procedures for teacher evaluation outlined in New Jersey’s TEACHNJ follow the same punitive

logic. The teacher evaluation system is set up to punish the individual despite any systemic or structural failings of the school and society. One of the biggest reforms provided with this law, which is casually referred to as a “tenure reform law,” is that teachers could no longer be guaranteed tenure if their students did not score high enough on standardized state tests. In this case, the punishment is an increasing loss of control by the teachers. First, teachers who are rated as “ineffective” or “partially effective” are put on a correction action plan (CAP). The legislation calls this “charges made against a school employee” as if they have committed a crime. The teacher has a hearing, and then that teacher’s tenure is revoked, removing their job security, one of the driving reasons that teachers used to go into the profession.

### ***Stakeholder Involvement***

In 2018, after Governor Murphy took office, New Jersey conducted an outreach effort to meet with stakeholders across the state as part of phase one of outreach to gain feedback on its assessment. Educators offered suggestions about the use of standardized test scores in evaluating teachers, particularly that it should only be used for specific groups of students, should be consistent across all teachers, and that the weight should be reduced or eliminated. The participants were not identified as either teachers, administrators, or another role. Eventually the state did reduce the use of standardized test scores to 5% and eliminated it during the two years disrupted by the start of the COVID-19 pandemic. The department of education also shortened state testing in grades 3-8 and high school by approximately 25% to allow for more classroom time in spring of 2019. “This was to increase instructional time, enable school districts to reallocate resources and to better support students’ academic needs,” (New Jersey Department of Education, 2021). The state also mentioned that New Jersey educators are part of the bias and sensitivity review process that the Department coordinates as a standard operating procedures for

all Statewide assessments in accordance with best practices in large-scale assessments (p. 81)  
The participation is noted, but the depth of the participation and impact is unclear.

In the 2011 pilot study conducted by researchers at Rutgers University for the State of New Jersey, survey results indicated that teachers in the study were not as positive as administrators regarding teacher evaluation. They agreed to the necessity and usefulness in evaluation to raise standards of practice, but “they disapproved of evaluations that primarily supported managerial decisions and aimed at meeting minimum standards” (Firestone et al., 2011, p.27). Most teachers in the study reported that the evaluation rubric created tensions among staff. The 2011 pilot study also identified concerns about inter-rater reliability: “...while many administrators see the new protocols as objective or increasing the objectivity of the evaluation, teachers disagree” (p. 30). Data from teacher evaluations is supposed to drive professional development, but the report states that usage was “limited and diffuse” (p.20). At the time of the pilot, some administrators felt they became aware of their own bias in the process, but the majority of administrators felt that the training did not help them understand potential biases in the way teachers are evaluated. Teachers felt the data collection on laptops distracted the supervisor during the formal observation to the point that they missed key details in the lesson. A consistent finding was that “teachers viewed the fairness, accuracy, and usefulness of the teacher evaluation program less positively than administrators and that they were less convinced of the adequacy of available evaluators” (p.46). The state’s response to this finding was to offer more training to teachers to give them a more positive perspective on the new framework.

In the 2011 pilot, the state responded to educator feedback regarding how to shape the implementation system, AchieveNJ. There were some procedural responses to feedback that

included delaying implementation, extending deadlines, adding more discussion, and conducting some additional research. The state also adjusted universal observation and training, created more training, reorganized their website, streamlined procedures, enhanced communication with districts, and expanded upon resources available to districts including templates and exemplar SGOs.

When educators wanted a “collaborative, transparent approach to adopting new evaluations,” the department responded with changes to the way they were conducting the pilot, not to AchieveNJ itself, by changing the format of their Evaluation Pilot Advisory Committee (EPAC) to include more collaboration. They also required districts to have District Evaluation Committees (DEAC) and School Improvement Panels (ScIP) with parents, board of education members, administrators, and teachers working to collaborate on teacher evaluation at the local level. DEACs oversee and guide the implementation of the local district policy with various stakeholders listed above, while ScIPs involve only the principal (or designee) and teachers. The DEACs’ functions are involved in implementing the system: coordinate planning efforts, clarifying existing rules, provide consistent vision in the teacher evaluation process across the district’s schools, and develop professional development plans based on evaluation data. The district power in the teacher evaluation process is limited, so while DEACs and ScIPs may discuss or process the issues in collaboration, it was never an intention of this aspect of the policy to move local knowledge back up to the state level and remains unclear how local educator input could be incorporated into any changes.

When district leaders and educators reported that they felt overwhelmed by the complex requirements of the new system, the state department responded with a few minor changes, such as streamlining communication and offering a review process for educators who were negatively

affected in the first year of AchieveNJ's implementation. Unfortunately, the changes were framed in ways that appeared more substantial than they really were. The state department reduced the weight of the standardized test score (SGP) from 30% to 10% in 2014. In Table 5, below, it appears this was due to feedback. In another official document, it was stated that this reduction was due to the administration of a new standardized test, called PARCC.

**Figure 5**

*Educator Feedback and State Response in Shaping AchieveNJ*

**Appendix D: Educator Feedback and State Response in Shaping AchieveNJ**

What the Department Heard	In Response, the Department...
<b>Pilot Years (2011-13)</b>	
Time constraints, heavy training load, balancing district activities provided significant challenges.	<ul style="list-style-type: none"> <li>• Delayed statewide implementation</li> <li>• Set capacity-building requirements for all districts</li> </ul>
At end of SY11-12, non-tested grades and subjects, and summative ratings had not been adequately addressed.	<ul style="list-style-type: none"> <li>• Researched these topics in expanded pilot</li> <li>• Discussed issues in EPAC meetings</li> <li>• Shared information in reports/communications</li> </ul>
Collaborative, transparent approach to adopting new evaluations was most valued.	<ul style="list-style-type: none"> <li>• Added time for collaboration in EPAC meetings</li> <li>• Required District Evaluation Advisory Committee (DEAC) and School Improvement Panel (SciP) groups statewide</li> </ul>
EPAC advised districts should choose from a wide variety of high quality observation instruments	<ul style="list-style-type: none"> <li>• Maintaining and updating state-approved list of teacher and principal practice instruments</li> </ul>
Process for ensuring number, accuracy, and fidelity of observations needed improvement.	<ul style="list-style-type: none"> <li>• Required training for all districts as part of capacity-building</li> <li>• Established universal observation requirements</li> </ul>
EPAC advised DOE to balance announced/unannounced observations, support new teachers with multiple observers and pre- and post-conferences, and incorporate double-scoring.	<ul style="list-style-type: none"> <li>• Built these parameters into universal observation and training requirements</li> </ul>
Districts did little to prepare to use observation data to make personnel decisions or to plan collective professional development.	<ul style="list-style-type: none"> <li>• Shared recommendations in final EPAC report</li> <li>• Linked PD to AchieveNJ in recent PDP templates</li> <li>• Continuing to develop relevant guidance</li> </ul>
Evaluation rubric should be simplified in first year of full implementation.	<ul style="list-style-type: none"> <li>• Included one practice instrument and two measures of student achievement (mSGP, SGO)</li> </ul>
State must acknowledge SGO learning curve.	<ul style="list-style-type: none"> <li>• Limited total SGOs to 2 at most</li> </ul>
Frequent and accurate two-way communication needed between DOE and educators.	<ul style="list-style-type: none"> <li>• Required DEAC and SciP groups statewide</li> <li>• Provided communications recommendations in Final EPAC Report and on website</li> <li>• Planning additional guidance</li> </ul>
Growing size of EPAC plus meeting structure/time constraints created frustrations among members.	<ul style="list-style-type: none"> <li>• Improved EPAC meeting structure/feedback mechanisms in second pilot year</li> <li>• Designed ANJAC to be smaller, more interactive</li> </ul>

What the Department Heard	In Response, the Department...
<b>Statewide Implementation Year 1 (2013-14)</b>	
State advisory committee should continue.	<ul style="list-style-type: none"> <li>• Launched ANJAC in November 2013</li> </ul>
To maximize DEAC potential, state should provide more guidance on their operation.	<ul style="list-style-type: none"> <li>• Solicited input from ANJAC</li> <li>• Launched DEAC Corner page of website</li> </ul>
In many cases, SGOs have been set with top-down, compliance-based processes.	<ul style="list-style-type: none"> <li>• Produced clarifying materials/workshops stating explicit intent that SGOs should be teacher-driven</li> </ul>
Educators need more information about high-quality SGO assessments and target setting.	<ul style="list-style-type: none"> <li>• Produced SGO 2.0 materials/workshops focusing specifically on these areas</li> </ul>
Educators need a wider variety of SGO examples.	<ul style="list-style-type: none"> <li>• Working to triple exemplars</li> </ul>
October 15 deadline for SGO approval is too tight.	<ul style="list-style-type: none"> <li>• Proposed to extend SGO deadline to October 31</li> </ul>
Teachers struggle with effective discussion and questioning techniques in the classroom.	<ul style="list-style-type: none"> <li>• Offering statewide workshops for teachers on this topic</li> </ul>
Information is not reaching all audiences who need it.	<ul style="list-style-type: none"> <li>• Increased scope of communications via county offices and superintendents, superintendent round tables, principal email newsletter, <i>The Bridge</i> newsletter</li> <li>• Tailored support for delivery through many channels (implementation managers, workshops, website, etc.)</li> <li>• Reorganized website to make resources more accessible</li> </ul>
District leaders and educators feel overwhelmed by new initiatives, too much information and are concerned about impact of new measures such as SGOs and PARCC on evaluations.	<ul style="list-style-type: none"> <li>• Initiated changes to streamline evaluation processes</li> <li>• Proposed review process for educators negatively impacted by SGO score alone for 2013-14</li> <li>• Proposed reduction in weight for PARCC component (mSGP) for 2014-15</li> <li>• Produced key points for principals and county offices</li> </ul>

The EPAC follows federal mandate that “a diverse and representing group of stakeholders” was required to be consulted in the states’ process to set their goals (citation). Additionally, a Committee of Practitioners is required by Section 1503(b) of ESSA. New Jersey representatives also met with a variety of organizations to advise them on their consolidated plan to implement ESSA in 2017 and established this committee to advise the agency in carrying out its responsibilities under Title 1: “The committee provides a field perspective regarding programs authorized under the law by identifying local implementation opportunities and challenges, discussing and providing advice on policy issues, recommending possible solutions for problems and identifying promising/evidence-based strategies for replication” (NJ Consolidated Plan, p.36).

Seven years after its original passage, the TEACHNJ Act was readopted by the New Jersey Department of Education with minor amendments in 2020. The *New Jersey Registrar* publishes all proposals for rule-making and public comments received. In 2020, when TEACHNJ was being considered for re-adoption, the New Jersey Department of Education received 252 comments from 2,124 educators and other members of the public in attendance, a much higher than average number of comments received during public hearings. Based on the minutes in the *Registrar*, the state replied to each comment, sometimes grouped in batches due to similar content, but no clear changes to the legislation were indicated as a result of these comments. Most comments were about standardized testing in some form, and as noted above, the weight of standardized test scores was eliminated or kept low at 5% since 2018. One is left to wonder if the larger number of comments and public pressure helped keep that weight low in the years that followed, something that is not fixed in the legislation but varies from year to year based on political climate.

Analysis of these documents revealed that the state responded to educator feedback in the following cursory ways: minor procedural adjustments, additional resources, and reducing time constraints. Still, the originally designed framework stayed intact. In public hearings, the state referenced compliance with federal rules as the reason behind components such as linking teacher evaluation to student outcomes. It was clear from the state's unresponsiveness to feedback that the public hearings were disingenuous, and the priority was compliance with the federal government's mandates.

Unfortunately, students' voices are not prioritized in the mandates either. Students are treated like objects in the teacher evaluation system. Following the logic of the federal mandates, students are framed as requiring fixing, with terms like "high-needs" and "disadvantaged" used



throughout New Jersey's official documents. This is in compliance with the federal government's framing of students. In the long list of organizations that were consulted and in the public comments received by the state educational agency throughout the implementation process, no stakeholders were identified as students.

### ***Illusion of Agency at the Local Level***

By studying these official documents, the ways in which power operates in the teacher evaluation system became apparent. "Policy documents have been critically described as a demonstration of state power, which do not include all relevant information and can be written in ways so as to mislead some groups of readers (Codd, 1998 as cited in Grant, 2019)." Policy reports contain key sections and jargon not immediately accessible to the average person, such as teachers, despite being embedded in a self-proclaimed democratic process.

Not only is the language inaccessible to teachers, but the way the documents work together and are stored by the state creates a challenging puzzle to assemble. When conducting this document search and analysis, I found reports and documents buried inside other documents. Additionally, there were a substantial number of dead hyperlinks on the state website for older yet essential documents, which I had to search for in a roundabout way. This process mirrors what I experienced initially when trying to look at local policy with references to state statutes that, in turn, referenced federal documents. It was insubstantial to consider local policies in isolation since they were connected in a larger web of teacher evaluation policy. One is left to wonder how teachers can participate in the teacher evaluation policy system with so many elements of it remaining inaccessible to them.

The state department asserts there are opportunities for agency down the hierarchy of government, but this is a farce. For example, according to the final report regarding

AchieveNJ's implementation, there is a claim that good teaching is defined locally and "many districts have taken this opportunity to redefine what good teaching looks like in their classrooms," (p.21). While at first read, it may seem like districts are engaged in a creative process to redefine good teaching, the policies provide specific guidance that defines effective and highly effective teaching. Therefore, it is not that the districts are redefining good teaching, but that they are adjusting their expectations of what good teaching should look like according to state definitions. Perhaps AchieveNJ emanates an illusion of power by the district in this area.

Additionally, there is an illusion of agency regarding the setting of student growth objectives (SGO) each year between teachers and administrators. As one component of a teacher's evaluation, the SGO is determined by the teacher and principal as a classroom-based objective they want the students in that class to achieve. At first look, this measure seems to foster teacher agency because it appears to give them an opportunity to set a learning goal, outside of test scores, that is meaningful for them. Yet, this does not operate outside of administrator approval and oversight, which changes the tone of that process. Since principals' evaluations are based on whether the students achieve these goals, they have a conflict of interest in supporting the teachers' agency regarding their own goal setting. They may be more apt to make a suggestion to teachers regarding what goals they should use or to use the process as an opportunity to get the grade level or school aligned on another strategic objective. Additionally, the supervisor oversight changes the tone of goal setting from something that is potentially intrinsically valuable to something equivalent to administrative paperwork. For that reason, there may be less thought and ownership over the student growth objective. So, going along with the administrator's goal is not only the easiest path in a busy PreK-12 teacher's world, but sometimes depending on how it is presented, it may seem like the only path. Teachers also may not have the

researcher or psychometrics background to develop it well and may end up writing out the goals in mathematical ways that are not actually aligned with the instructional goals they had in mind. All these factors and more make the setting of student growth objectives less empowering.

### **Conclusion**

After analyzing the discourses at play in these documents, it appears that the explicitly stated goals (to improve teacher quality in order to raise students' scores on state tests) at the federal and state level are not telling the whole story. In fact, state test scores still remain low despite following these mandates for the last two decades (Tucker, 2014). This document analysis has revealed the underlying political discourses targeting teachers.

It also revealed that teachers are effectively *prevented* from participating in policy mandates in meaningful ways and made me aware that administrators and state level officials may feel the same constraints on their participation. Even so, all actors that comply perpetuate the way the system works, and it continues to operate.

As Grant (2019) explains, "official documents are often ideally suited to qualitative research" (p.62), so exploring policy documents in alignment with state and federal statutes makes sense in this study. "Documents provide the researcher with information about things that cannot otherwise be observed or about which the researcher was unaware" (Patton, 2002). These documents can be examined in relations to the teacher inquiry group or interviews but provide unique insight about the larger federal and state policy mandates that could not be gleaned from the other empirical data. Additionally, exploration of the larger political systems at play in teacher evaluation helped me contrast teachers' experiences from the intention of the policies in the chapters that follow. I gained insight regarding the larger system beyond what teachers encounter each year in their evaluation. I also learned about the technical documentation of teacher evaluation as I dug deeper into the policies that govern teachers' daily experiences in the

classroom and schools. Knowing this information helps me tie in teachers' thoughts to a larger system and understand the social function of these documents in the chapters that follow.

## V - FINDINGS

In the previous chapter, I presented a narrative about the discourses underlying the official text of teacher evaluation mandates. I also discussed the apparent theory of action based on market fundamentalism and global competitiveness manifest in these mandates. Federal mandates operated under the premise that improving teacher quality closes the achievement gap, but, in reality, they acted as a punitive measure to punish teachers. I also argued that these official documents are inaccessible to teachers, unresponsive to public comment, and contain standards that are federally, not locally, defined. In this chapter, I present findings regarding how teachers in the study respond to teacher evaluation policy mandates.

The major finding is that teachers in the study perceive the enacted teacher evaluation system in this school district to be harmful without much positive impact on professional development. Teachers in this study feel under attack and powerless in the face of teacher evaluation policy mandates. As discussed previously, teachers' evaluations are comprised of a formal observation of their teaching practice combined with data from student assessments (New Jersey Department of Education, 2017). The formal observation component of the system was more evocative than the other components for teachers, and their memorable recollections of these episodes inspired negative feelings. Teachers in the study challenged the notion that formal observations present a faithful rendering; it is just not enough to truly judge their teaching and it is generally unfair. They felt the teacher evaluation system contained an oversimplified view of children that was incompatible with their own held beliefs. Teachers in the study also were concerned with the negative impact of student assessments, especially state mandated tests, on students. In my analysis of their responses to teacher evaluation policy, I find that the teacher

evaluation mandates influence the teacher-administrator dynamic in ways that do not support student learning but, rather, the mandates serve to instruct teachers how to navigate the process. The teachers perceive the administrators to be biased and generally unhelpful in their post-observation feedback. However, teachers are compassionate about the constraints administrators face in this system as well. In the end, teachers in the study support each other, remain optimistic and dream up suggestions that might help improve the policies and procedures in their school district.

As these findings have the ability to inform teacher advocacy efforts, I am guided by Brittany's insight that an effort to change teacher evaluation policy would have to be "more of a collaborative effort to come up with a solution rather than a blame game." Although these findings present many critiques of the teacher evaluation mandates, my intention is not to blame any individuals for the failings of the larger system. I approach these findings with the hope that they can help other teachers and administrators work toward repairing the harm done by formal observation and misuse of standardized test data and are replaced by more liberatory practices. As mentioned in Chapter 3, these findings are contained in a mostly white, mixed rural-suburban public-school district and may or may not be replicated, with local variations, in other districts. The strength of this qualitative study is that it provides original insights, grounded in the co-construction of knowledge about teacher evaluation, among practicing teachers who are at different stages in their careers. These in-depth insights are not meant to be generalized, but rather inform teacher advocacy efforts with implications for professional development and future research as well.

### **Under Attack and Powerless**

The teachers in this study were aware of deeper discourses in the teacher evaluation policy mandates. Rose, Dorothy, Teresa, Nicole, and Lauren uncovered that the purpose of the policy mandate was to monitor teachers and put them on notice for possible elimination. To them, the policy mandates seemed to be indicative of a lack of value for teachers. Naomi and Brittany also questioned whether supervisors trusted them to do their best teaching. Teachers in the study felt under attack and powerless in the face of teacher evaluation mandates.

### ***Teachers Under Attack***

During the teacher inquiry group discussions, Rose, Dorothy, Teresa, Nicole, and Lauren interrogated the explicitly stated purpose of teacher evaluation to improve teaching and students' test scores. Rose provocatively said that teacher evaluation is "to improve your teaching. That really is the goal, but *is* that really the goal?" She identified the tension between what was stated in the official documents versus what she encountered in her lived experience, and was hinting at that apparent dissonance. She didn't answer her own question, but she saw there was something more going on.

The discussion during the inquiry group uncovered the ways teacher evaluation functioned in their lived experience to punish teachers and put them on notice for possible elimination. Rose shared about a change in teacher evaluation that she noticed during the span of her career which started when those federal policy mandates were passed. Rose said, "It used to be more to help you as a teacher than punitive," highlighting the increasing surveillance and potential for punishment associated with teacher evaluation over the course of her career. Rose did not have any awareness of the state or federal critique of the punitive measures from No Child Left Behind (NCLB) and requirements set forth by Race to the Top (RTTT) described in more depth in Chapter 4, yet she noticed something similar in her experience.

Dorothy and Nicole recognized the role of teacher evaluation in eliminating certain teachers. Dorothy also questioned the purpose of teacher evaluation, “Does saying that you’re proficient, or not proficient, it’s mostly for personnel decisions, isn’t it? Or what else? Truly it’s not helping us as teachers. It’s for decision making in the district, right?” Nicole agreed and stated “So, I find policy is the protection of the ruling whatever that is.” She described teacher evaluation policy as “the school district’s insurance policy, basically saying, ‘Well you weren’t effective, so now we can move you, cite you, get rid of you.’” Nicole identified how districts use the policy to protect the district not to advance teachers’ professional development.

Lauren said: “there’s definitely been teachers in the past who had fine evaluations and then all of the sudden in the last year they haven’t,” suggesting that the worsening ratings meant these teachers are under scrutiny by the administration. It wasn’t that they would be removed immediately, but getting lower scores on their evaluation was a signal that they were being watched. I shared with the group how I noticed a downward trend in my three observation ratings during my first year teaching. I remarked to the group, “It was like they were making a case to not renew me.” My first two observations looked fine on paper, but then the last one showed poor marks. It became apparent to me that my principal was building evidence to not renew me. Another supervisor observed me for my final observation and gave me exemplary marks. In that example, teacher evaluation was being used in a specific way to eliminate me based on my principal’s preferences, until another administrator came along and had a vastly different and positive opinion of my teaching. I held the conception that this saved my job.

All of the teachers in the study had tenure, so they were afforded job protections despite poor ratings on teacher evaluations. It wasn’t the threat of poor ratings that worried these teachers at this point in their career. To them, the teacher evaluation system further exemplified



society's lack of value and trust for teachers as professionals who knew how to do their jobs. Teresa stated in her interview, "We're not, you know, valued at all because you know everything keeps getting worse for us instead of better." Teresa described how additional federal mandates were placing more and more requirements on teachers that got in the way of teaching children. She said "They keep putting these things in place that they're trying to make us accountable but they don't understand our profession and what we're doing. And they have no idea what's going in the classroom, so it doesn't make sense." The mandates showed a lack of understanding of her work, and they did not put forth tactics that would improve her teaching.

While the teachers did not have direct experience with policymakers, Naomi and Brittany questioned whether those who implemented the teacher evaluation protocols, i.e., supervisors, trusted them to do her best teaching. Naomi pointed out that particular supervisors trusted teachers more than others. She remarked that some supervisors "said this is what you have to do and other supervisors, like one currently, let us use our best judgement. This is what we have. This is the program, but if it doesn't work for your class, use your best judgement, like if you have to supplement." Brittany knew who Naomi was referring to and remarked that this supervisor was recently a teacher, so she "makes judgment calls as a teacher rather than a supervisor." Brittany was saying that this supervisor's long history of teaching herself caused her to trust teachers more than a supervisor who did not have that teaching background; for example, someone who did not teach for a long time or who did not teach recently. In her interview, Brittany said we needed to have a "trust that people are self-motivated." She felt the majority of teachers in the district were "highly motivated, highly reflective and just want to get better." However, supervisors did not always act in ways that reflected that strengths-based perspective of teachers. Administrators designed teachers' schedules and professional development activities

in ways that did not honor how the teachers felt they needed to spend their time to improve their practice.

During the last teacher inquiry group session, Tracy said, “I think administrators fear that if they give us that time, it’s not going to be productive, but really it is one of our most productive times.” Even though the district had some power to show value for the teachers’ expertise, the district leaders did not cut back on the time-consuming tasks they assigned teachers. Administrators hesitated to give teachers extended blocks of “free” time. In her interview, Brittany suggested “if every district gave teachers reflection time in the day, they could use that time to talk to each other or talk to coaches, and not fill it with all the extra baloney they have to do. That would be amazing.” Teachers’ time was being used by administrators to accomplish tasks aligned to a strategic vision into which the teachers had no input.

### ***Powerless***

Teachers in the study felt powerless when it came to certain decisions in their district and also in their ability to impact the federal policy. They advocated for their students to get more resources, but the district could not always provide them. Teachers also felt substantial time constraints and a lack of ability to pursue personal interests of growth, including conduct teacher research. Teachers in the study discussed that they could attend local Board of Education meetings as a group to try to create change, but ultimately, they settled on the idea that voting in elections was their major vehicle to creating change.

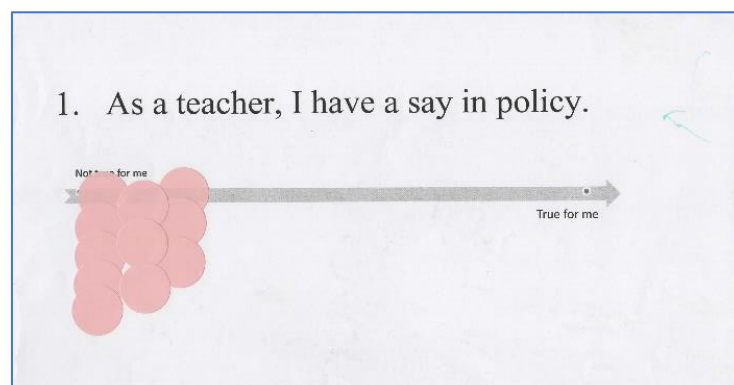
Teachers were advocates for their students to receive the resources they felt the students deserved yet were often left feeling powerless since their requests for support were ignored. These teachers shared stories of attempts to get students tested for additional support within the

district and being denied even the initial testing that would determine if students qualify for special education resources. Melissa said, “And you want to help the kids that want to learn but you can’t.” Naomi commented that not being able to reach a student because there was not enough support caused her anxiety. No matter what reason was presented, teachers found the lack of support upsetting. They felt caught in the administrative bureaucracy with no power to get the help they felt their students needed.

Similar to the feelings of powerlessness in obtaining help for their students, teachers in the study felt an inability to have a direct impact or influence on policy. During the second inquiry session, teacher reactions to the statement, “As a teacher, I have a say in policy” prompted reflection on their role in policy (see *Figure 6*). Kimberly suggested they could go to a Board of Education meeting, but several of the other teachers disagreed with that idea. Brittany responded “I don’t think that makes a difference. I think they listen to parents more than teachers.” They felt showing up at a Board meeting would be ineffective due to the perception that the Board does not listen or act on teacher statements in that forum. The group decided that attending a Board of Education meeting would not be an effective method or means to influence teacher evaluation policy.

### **Figure 6**

*A Photograph of a Consensogram Constructed by Teachers*



In her interview, Teresa spoke about a time when her teacher's union organized teachers to show up and speak at a Board meeting to support a contract settlement. This was an inspiring moment for her, as it showed the potential for mobilization efforts in affecting change; in this case related to the need for salary increases to maintain teacher health benefits. But she felt influencing teacher evaluation policy was different from settling a contract. She said administrators didn't have the power to change the policy, and that teachers would have to go "higher" than administrators to make a change. She felt it could be addressed if teachers and administrators worked together to speak up, but that "getting everyone on board to do that is the tricky thing probably." Teresa was hinting at a larger effort that involved teachers and administrators to come together to advocate for teacher evaluation policy changes at a higher level. Teresa said "When the evaluation first came out, maybe there were actions being taken. I think it has just died down and everyone is accepting it...I know teachers were upset about it but I don't know necessarily. I didn't see really anyone taking action unless they were going to go support the NJEA, but I don't really know of anyone who did that." Teresa suspected there was advocacy around the issue, but she felt removed from those efforts and unmotivated to pursue action alone.

Teachers in the study also said the three case examples of teachers impacting policy (see Chapter 3) shared with them toward the end of the inquiry group work, and again during individual interviews, did not resonate with them. Brittany said, "it's just a nice reminder that teachers can make change. I think sometimes we're kind of following rules and most teachers are rule followers so, it's not necessarily making waves. I think you're afraid to speak up and say anything that's not positive." Teachers also talked about the limits of time in being able to participate in teacher research despite feeling it was important. Teresa shared that she thought

teacher research was a good idea, but she said “time is the factor.” She said, “if we can spend our time doing more things like that instead of what we already are doing with this, you know, recording of data to prove that our students went up three levels or whatever silly things like that. But, yes, I think that is important and helpful.” Their engagement in the discussion showed that the topic of teacher evaluation was meaningful. During follow-up interviews, they also expressed that being together in community across schools was the most meaningful aspect of their participation in the study. However, time was a limiting factor in their own ability to explore or conduct teacher research to create change.

In her interview, Kimberly settled on just accepting the status quo. She said, “I feel like it is what it is in our district, you know. Just kind of go with it.” She did not feel the need to take action or investigate the issue further. If a group of teachers decided to take action, she said she would need to know why before joining them, but would consider it as long as it was comprised of a large number of people already organized into action. During the inquiry group, Dorothy was skeptical about how teachers could influence policy. She commented that policy is “all predetermined by other interest groups. I don’t know. It’s the money, it’s the policy, it’s the federal government...” The teachers felt removed from the policymaking process—a finding that I reported from my document analysis in Chapter 4 as well

Toward the end of the discussion, Rose shared the idea that voting was the best chance to influence policy. When asked if the teachers in the study felt like they had a say in policy, Rose replied, “The say, although small, is your vote. And then it’s out of your hands.” The group agreed on this approach. They felt the best pathway to improvement in teacher evaluation was to elect officials who would make better policy decisions that supported teachers. Although they had some examples of effective civic engagement involving teachers’ unions organizing and

groups of teachers researching solutions which they published, they settled on civic engagement through voting. Overall, when asked how teachers could affect policy, teachers in the study doubted themselves and their ability to participate in policy but settled on the idea that voting was their main vehicle to change.

### **Responses to Formal Observations**

As discussed at length in Chapter 2, the teacher evaluation system in New Jersey is comprised of formal observation and measures of student growth as determined by pre- and post-assessments. For teachers of core subjects, such as language arts, mathematics, and science, the students' growth on state tests from the previous year comprise a portion of their total evaluation. For other teachers, the scores are based on district-based assessments or sometimes those that are administrator or teacher designed.

The formal observation component was significant to teachers in the study. This was to the point that teachers saw the term “teacher evaluation” as synonymous with formal observation, even though the official policy described it as just one component of a reformed teacher evaluation system. Historically, formal observation was the main component in teacher evaluation mandates before reformed federal mandates tied teachers' evaluation to student growth objectives and student performance on standardized tests. As stated in Chapter 4, measures of student growth were not formally factored into New Jersey teacher evaluations until after the state won its Race to the Top waiver in 2011. But, somehow over a decade later, the belief the formal observations were the sole measure of teacher evaluation persisted among teachers. Since this component was so significant, it seems fitting to spend time in the findings exploring their responses to the formal observation further. During the first inquiry session, teachers shared recollections of their experiences with formal observations in a storytelling

circle. I found that the shared stories about formal observations reflected stress and anxiety without much space for teachers' perspectives.

### ***Negative Emotional Responses to Formal Observation***

The formal observation component of the teacher evaluation system was more emotionally evocative than the other components for teachers in the study. And, the most memorable episodes of formal observation for teachers in the study were not positive. The teachers' shared recollections of formal observation that included some sort of negative emotional response to some aspect of the observation process. Some teachers felt stress or worry in the time leading up to a scheduled visit, during an unannounced observation (a "pop in") or in response to post-observation feedback from supervisors. The direct, firsthand involvement of teachers in this study provides evidence that the formal observation process was experienced as emotionally harmful to the teachers.

For example, Dorothy shared the negative feelings the formal observation process inspired during the first session of the inquiry group. She was teaching at a public school for the first time after teaching at private schools that did not conduct teacher evaluations or any type of report on one's teaching. In an interview, she explained the conditions at the private school further and commented "As long as you were willing to continue working, you were working. I guess until something bad happened." Her employment contract was at will rather than the product of collective bargaining by a teacher's association. Yet, she didn't feel the high stakes pressure she would come to experience in a public school district through the formal observation process. The experience of being observed was brand new, and she felt intimidated, nervous and describe it as "stressful" going into the lesson. She shared with the principal that she had several students with misbehaviors. Dorothy explained that the principal said "my big focus was just

trying to do classroom management.” The principal said she was less focused on the lesson itself. Dorothy repeatedly emphasized how “distressing” and “worried” she felt going into the lesson. She said “I did it much more whole class and I didn’t give much leeway to the students, because I was worried that it would just fall apart.” She went on “It bothered me that I was really not getting help for some of my students and expected to handle it myself.” Dorothy said she did not feel this experience improved her practice, and she felt a lack of support in helping the students who needed it. Her takeaway was “You kind of make it through.” The observation caused negative feelings leading up to it, and Dorothy was left feel unsupported afterward.

Naomi also shared negative feelings inspired by her recollection of teaching verbs and nouns during her early career. This was before the current teacher evaluation mandates. She had met with her mentor at the time to plan and decided to have students work in pairs to read a poem to ignite the lesson. After a supervisor observed her, she received what she perceived to be a very negative write-up of the lesson that focused on student misbehavior and urged her to be more of a “team player” by talking with the other teachers on her grade level. Naomi reflected on the experience. She described her feelings as “grateful” to learn about misbehaviors she had missed during the lesson but “extremely upset” and “offended” because she had actually planned with her team. For the next lesson, Naomi decided to present a very teacher-centered science lesson, and the supervisor loved it. The teacher learned about that administrator’s bias for teacher-centered lessons and was able to cater to her preferences in subsequent lessons in order to earn higher marks. She said “I also felt that she was looking for a very specific type of lesson, that was her personality. And if they had been better behaved, I’m not sure she would have loved that first lesson anyway.” Naomi learned if she wanted to be successful in this particular administrator’s evaluation of her teaching, then she should structure the lesson in a teacher-



centered way. She said the process did not improve her practice. Through this experience, Naomi also endured what she perceived as insults to her preparation for the lesson. Naomi said, “Anxiety is no joke for some with it, and these evaluations cause unnecessary stress.” Her responses show that she learned how to navigate the system by designing lessons aligned to the supervisor’s preferences in order to receive a better evaluation but endured a negative emotional response in the process.

Nicole also felt negative feelings as a result her interaction with her supervisor after a formal observation. She shared a recollection of being observed in 1999, a time that was before the current teacher evaluation system. She did not have a budget to purchase supplies for her room, they were not provided and the language arts program was outdated. She described the room as bare, “There was not a single game, book or toy on the shelf.” The students in her class had multiple disabilities, and her task for the day was to teach them the letter ‘x’. As a first-year teacher, she tried to make the lesson exciting by creating and buying “all kinds of hands-on activities” for the students to do. The supervisor stayed for an hour and remarked at the end, “Good job for such dry subject matter.” The teacher shared that, after that comment, she felt “completely deflated and did not know where to go next. I felt like I did a terrible job and what should I have done to make it better because I had to cover the material.” The district did not provide the tools she needed to teach, including teaching tools, a budget for supplies, and an up-to-date language arts program. The supervisor’s comments placed all responsibility for the shortcomings in the lesson on the teacher and overlooked the structural problems which, as a supervisor, he may have been in a position to address or ameliorate. To Nicole, her supervisor’s comment showed his lack of understanding for her students, her instructional practice and her feelings of competence in her work.

Kimberly shared that her first formal observation of her career did not go well. This took place after the current teacher evaluation system was introduced. She focused on classroom management issues, like materials spilling and general chaos. A student intentionally flipped a cup of water. There was water all over her classroom floor. She said “It was craziness. I was embarrassed. Sweating profusely. I was so scared. And it was really fast, like September or the first week in October...And it was unannounced.” Kimberly didn’t have the lead time to plan this lesson in a certain way, such as a more teacher-centered way to avoid accidents or to match the supervisor’s preferences. Due to the surprise element, she did not have to endure the stress that others felt before the lesson. She said, due to the lesson being unannounced during her first year teaching, she experienced embarrassment, fear, the physiological symptom of sweating during the lesson.

The vignettes above illustrate how negative the formal observation process was for the teachers in the study, whether they were announced or unannounced to teachers in advance. There was no evidence that teachers experienced an interruption or a change in thinking about formal observation after the current teacher evaluation system was introduced. The teachers commented about observations throughout their careers as if there were no policy change. The formal observations remained a constant source of harm despite federal reforms. As detailed above, Dorothy felt distress and worrying leading up to the announced observation, and she felt unsupported afterward. Naomi and Nicole both felt prepared going into their announced formal observations, yet the supervisors’ remarks afterward made them feel poorly. Supervisors’ post-observation comments were perceived as offensive for Naomi and deflating for Nicole. Kimberly’s unannounced formal observation caused her to have a physiological reaction which included sweating profusely and feeling embarrassed and scared. I selected these vignettes to

paint of picture of how teachers in the study perceived formal observations in ways that inspired negative feelings. The snapshots convey real emotional harm done to teachers during this process.

### ***An Unfair Process***

Teachers in the study perceived the formal observation as unfair. They challenged the notion that formal observations are a faithful rendering of their teaching. They said the teacher evaluation system did not account for different roles and curricular programs different teachers enacted. They also pointed out how administrator subjectivity made the teacher evaluation system unfair. Formal observations also inspired negative feelings for participants in the study, which are illustrated through their vignettes presented in this section.

One critique of formal observations was that they were unrealistic, rather than a true realistic description. Brittany said, “I practiced for weeks, exactly what I was going to say and that’s not realistic, it’s absolutely not realistic.” Kimberly defined teacher evaluation as “watching a teacher on their best behavior, cause that’s what it is usually.” She also said it was like “Ok, here’s the perfect little lesson. Everything’s buttoned up and the kids are angels.” If teachers were aware of a certain supervisor’s proclivity, they could design their next “show” to appeal to these preferences. However, it didn’t seem that was always a guarantee for high marks. In working with the teachers, I found that other factors, external to the incidence of formal observation, were at play in how an administrator evaluated them.

To Kimberly, Brittany and Teresa, the administrator’s evaluation of teachers was ongoing, not just something that took place during an observation. Kimberly talked about how the observation is just a formality in her interview, “It’s a small sliver of what a teacher does.

And it's the teacher being on their best behavior. I don't know how accurate it is. I mean, they get a good idea. Administrators know the people that they work with, the teachers in their building. It's more just a formality of securing an opinion about someone." She had the sense that the administrator's judgement of the teacher is crafted based on the administrator's work with the teacher each day outside of the formal observation process. Teresa suggested in her interview that only one formal observation would suffice for teachers who "haven't had any issues in the past." She said "if you're a tenured teacher that, you know, hasn't had any issues...I don't see the point of doing that as many times." Brittany contrasted her experiences with administrators who knew her well versus those who did not. She valued the observation feedback from the former, because she felt it wasn't a show. She said of the latter, that "they didn't know me personally, they didn't know my relationship with the kids, they didn't know my relationship with parents, they didn't know the makeup of the classroom." To Kimberly, the administrator already had an established view of how effective the teacher was and the observation just codified it on paper. In that case, the administrator judgement of the teacher is based on some other ongoing, unstated measures. These measures may or may not match the criteria put forth by the district-approved rubric. The evaluation is not even based on formal observation visit but, perhaps, ongoing, and actually taking place in another space over time outside the direct classroom observation.

Teachers in the study offered the critique that the teacher evaluation system was not differentiated and did not align with the various curricular approaches they were supposed to use depending on their job assignment. Kimberly pointed out that the work of specialists who teach subjects like physical education, art, music and so on, is not reflected in the Danielson (2008) rubric. Nicole also mentioned that inquiry-based approaches in science and math were not

captured by the models of excellence in the rubric either. Naomi wondered, “Our evaluation has a big area for questioning. So, if you have a lesson where the students aren’t supposed to be questioning, how does that line up?” Another issue was different class makeup each year. Dorothy said, “When you really put your heart into it, you end up getting a more difficult group every year and there are other teachers who...only get the children who sit still with their hands crossed. So how do you compare apples to oranges?” Dorothy was suggesting that teachers who do well end up with harder or more challenging students the next year, which might result in a poorer evaluation.

Teachers believed administrators applied a singular evaluation model that failed to take into account different teacher roles and responsibilities within a school district, not to mention comparison to others schools within the district or other districts each with locally adopted curriculum. Danielle said, “...The state wants to see continual improvement of instruction but we’re using different curriculums so are we all truly being graded on the same measures? Scales? Playing field? Because someone may have a basal reader where every day is scripted out and all you have to do is open the book and read it out to them. And then there’s another program where you create the entire lesson off an idea..., so I don’t know.” This teacher also pointed out that different districts have different tools and resources at their disposal. As explained previously, the state uses the same formula to calculate a teacher’s final score comprised of their formal observation and evidence of student growth determined by local assessments and/or state assessments. The teachers do not choose the curriculum, as that is the job of the superintendent and the Board of Education. So, teachers are at the whim of whatever tools are approved locally, whether they reflect the most recent research in teaching children or not. Danielle wondered how

the state could compare educators across districts when there is so much variability in their teaching assignments and the tools they are given to teach.

Teachers in the study also pointed out the subjectivity of administration in the ratings received during the teacher evaluation process. This challenge to inter-rater reliability matches the results from the 2011 New Jersey pilot study reported in Chapter 3. Teresa said, “you can do the same lesson and it can go the same way. If you have two different administrators you might get two totally different scores.” After reflecting on the terms that stood out to the group in the first two sessions, Nicole said that the term “subjective” was the biggest for her. She said, “I mean evaluations are truly subjective compared to whatever that particular administrator wants. You and I are going to see the same lesson and think differently about whether it’s the person and the delivery or if it’s about the lesson, looking at different things and for different things. Just depending on our beliefs.” Naomi, Teresa and Nicole suggested that administrators’ preferences were big influences on the outcome of evaluations.

These teachers knew that objectivity was not achievable through the current system of teacher evaluation. However, subjectivity was not something administrators were transparent about in their evaluation work. There was no acknowledgement of the different administrators’ biases in their evaluations, which could have buffered teachers’ feelings that the evaluation system was unfair due to subjectivity. Administrator bias will be discussed further in this chapter.

### **Responses to Other Aspects of the Teacher Evaluation System**

Administrators evaluate all teachers in New Jersey based on formal observations and student achievement of academic growth goals. Some teachers, who teach what the state considers to be core subjects like Mathematics, Language Arts and Science, face evaluation

linked to students' growth on state tests. In this section, I explain findings that teachers in the study who taught core subjects were concerned with the reliance on standardized test scores in the evaluation system. Their comments revealed that while they did not understand how the test scores specifically impacted them, they were concerned by the relative weight of the test scores as part of their student outcome evaluations. They also felt children were treated like numbers in the teacher evaluation system. They questioned the basis for the tests, including whether the content, based on state standards, was developmentally appropriate. which was based on state standards. They also did not like how the standardized tests required time away from their regular instruction and noted feeling rushed in covering material as a result.

### ***Oversimplified View of Children***

Another critique of the teacher evaluation system was that it treated children in a way that was inconsistent with the teachers' beliefs about children. The longer teachers in the study had taught, the more they saw value in connecting with children and responding to the complexity of children in their work. Brittany said "Getting a child to love to learn is the most important thing." She also remarked that a central part of her professional identity was feeling like "I put the child first." The teachers in the study viewed children holistically inside the classroom. Teachers embraces that the whole of the child's life impacted whether or not they were successful in school. Brittany commented "There's so much more to it than they come in, we teach them multiplication, and then they leave. So, I think understanding not only that they are human but there's other facets to the child." Brittany, Teresa, Dorothy, Nicole, Melissa, and Naomi shared that they used to be hyper-focused on academics, but as time went on, they understood the child as "a whole child.". Several teachers mentioned that as their career progressed, they changed their approaches, and went from prioritizing the delivery of a predetermined curriculum to a more responsive approach that would help students better retain knowledge. The end goal

remained focused on student learning, but their understanding of children as agents in that process seemed to increase as time went on in their career. This view of children influenced their outlook on the use of student growth goals in teacher evaluation.

Teachers discussed how students lead complex lives that were not captured by the contemporary teacher evaluation and associated testing systems. They discussed how more students seem to exhibit more issues in school, like anxiety. Naomi said, “the students have more complicated home lives and social issues than when I first started teaching. I don’t think they were as apparent or occurring as frequently, at least in this district.” Brittany shared that students show more anxiety “to the point that it’s debilitating.” As a result, she feels required to prioritize instruction each day in ways that can mitigate or avoid anxiety. Brittany explained “you get this physical empathy that eats you up inside because you have very little control when they leave your classroom. And, that is torture in a way. It makes you a great teacher, but it also physically screws you up.” Melissa agreed that her current students had more complex home lives than previously. She said “I also think policymakers should take a look at the group of students that are going through now. Because, the kids that are going through now are completely different than the kids that were going through 10, 20 years ago...and it’s a huge difference. And I know that it takes years and years to write policy but it has to be fluid.” Naomi stated that she was asked by an administrator to offer a second or third chance to those students who did not perform well on the local benchmark assessment. Naomi said that it was “nice to have [the extra time], but there are always going to be students who can’t jump over that bar.” She was doubtful that those students would ever perform to the same level as other students due to all of the issues they had at home. Brittany shared, “Like, you know sometimes a student needs more than you can offer. You’re not able to get them the help that they need. Or things are happening at home.

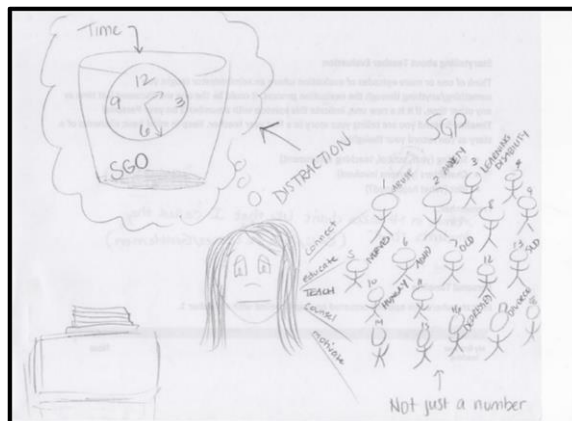


There's no asterisk saying this child is in an abusive family or that this child's parents are in jail." Dorothy said, "They want everything to be standardized and you know." She pointed out that there is an increase in families living in nearby hotels and wondered how those children could be expected to perform the same as students who had more financial resources. She wondered, "Like, why is there a need for that... trying to pretend that everyone is the same?" These teachers recognized the out-of-school factors in a child's life, something that was not captured by the federal mandate.

The teachers in the study expressed that the increase in state-mandated content standards, state testing and district curricula offered a view of children that was incompatible with their own personal view, and negatively impacted their ability to reach children. Many times, teachers shared that children were treated as only numbers in the teacher evaluation system, and the system did not embrace their full humanity (see Figure 5.2). Dorothy explained "Sometimes there are so many factors that go into a person, not just their test taking abilities or knowledge. There's a lot more there." Dorothy shared, "Sometimes things seem to be completely overhauled rather than allowing the teacher to use her professionalism to utilize things that are good teaching and best practices." Naomi described the treatment of children by teacher evaluation as "unsettling" since all children have been treated the same in this system.

**Figure 7**

*Not Just a Number*



### *Concerns about the Impact of Standardized Testing*

Teachers were also concerned about the impact of state-required testing on students. Melissa said “[Assessments] put unnecessary stress on the kids, because if you have to give a running record several times, or if you give the math multiplication every week or month. I think... I would not want to come to school that day.” The frequency of testing was also viewed as problematic. Teachers in the study viewed their own feelings about teacher evaluation as similar to the impact of state testing on the children; a required exercise that had negative repercussions. Naomi looked for a way to lessen the force. She wished observations would be reduced to once per year, as they were in the past, due to anxiety around the process. Teresa had also made this suggestion in her interview with me.

Just as time was a limiting factor in participation in teacher research, the topic of time came up frequently as an issue related to standardized testing constraints. The teacher evaluation system did not mention the workload explicitly, but clearly the expectations associated with the description of a successful teacher would be very time consuming, when combined with the numerous tasks already required of an elementary school teacher. The demands of elementary school teaching require practitioners to remain conscious of time when planning lessons, delivering lessons within the allotted time, and keeping up with administrative tasks on a daily basis. Their pacing in units must be speedy to ensure adequate coverage before the quarterly district-level benchmark assessments and the annual state standardized tests in May. The teacher evaluation mandates added additional constraints on teachers, since student growth is measured for only part of the year, sometimes terminating as early as March. Brittany expressed that she felt like she had to rush her teaching. She said, “You have to teach a whole year of stuff in eight months.” Just as teachers were pressured by administrators, they also felt pressure to go faster in their instruction in order to set their students up for test success.

Teachers made comments that questioned the rationale behind the state standards for instruction. New Jersey Learning Standards affect their teacher evaluations, since those standards drive many of the student learning goals on which teachers are assessed. Teachers in the study suggested that the standards were too high, leading to problematic expectations and results. Brittany explained that the standards were “developmentally inappropriate” for the students. She said her students were showing “no retention, because so much is being crammed into their little brains before they are ready for it.” Dorothy failed to see the rationale behind the standards, believing they had “no rhyme or reason.” Naomi said “But they need to look at the standards also. I think that does tie in to evaluation... You can have the best teacher in the world with high expectations but you can’t force a child that’s not ready.” Rose suggested the bar had been raised too high, with her students over the past few years unable to perform at the same level as those in the past, before the state standards and aligned tests were changed. Since teacher evaluation is partly based on student growth on state tests, challenging that premise disrupts the validity of teacher evaluation.

### **The Teacher-Administrator Dynamic**

Having administrators evaluate teachers creates a specific dynamic that could be problematic. Administrators judge teachers’ with possible punitive consequences. Administrators also rely on teachers’ scores to inform their own evaluation rubric. This causes administrators to micromanage or step in to intervene when a child is not producing the right outcomes. Teachers recognized that the administrators were under constraints to comply to board-adopted curriculum over the teacher’s own judgement. This compliance to the board above all else compounds the issue of trust mentioned earlier and feeds into larger discourse in society regarding the mistrust of teachers. Teachers also felt that the feedback they received was shallow. It is unclear if this is due to the brevity of the observation, lack of supervisor training or constraints the administrators

face in their own responsibilities. However, this problem contributed to the teachers' negative view of formal observation.

### ***Administrator Bias***

Teachers in the study suspected one constraint on administrators may be the pressure to adhere to a program. In the first session of the inquiry group, Nicole had shared that her supervisor referenced the current curriculum and told her that her math warm-up was too demanding, as her students would not be able to change their thinking that quickly. However, the next year, the supervisor's department began a new math program that actually utilized the approach that Nicole was using in her warm-up. The supervisor came back and apologized and said "I guess you were on the right track." Their supervision of teachers was tied to the current departmental program and when that changed, the administrator had a pathway to change their thinking as well. While the administrator was constrained by the program, Nicole felt like she was not as tied to the program as the supervisor was. She had enough teaching expertise to supplement the curriculum in original ways that were responsive to her students. Brittany questioned whether this was due to an attempt at upholding program fidelity and also wondered how mistrust of teachers fueled this behavior.

Nicole critiqued administrators' bias regarding students with disabilities as well. She said, "If you have an administrator that sees this class has a particular student or small group of children that are significantly disabled in behavioral or emotional ways, then they will mark that and you will get a great score. But, if you have another administrator who sees it as a behavior problem or a management problem, then you are going to be marked low." Nicole's recollection of her formal observation involving teaching the letter 'x' early in the chapter supported this perception of students as well. Her supervisor did not recognize the challenges associated with

teaching students with significant disabilities, and his deflating comments after the lesson reflected that.

Teachers also felt the weight of administrators' evaluations. It seemed to cause administrators to pressure teachers somehow to get their students to perform better on district and state assessments at almost any cost. As discussed previously, teachers execute Student Growth Objectives (SGOs) each year as part of their teacher evaluation. In this SGO, students take designated pre- and post-assessments to measure their growth. The more significant the student's growth, the higher the teacher's score. Teachers' high SGO scores boost administrators' scores as well. Teresa knew of administrators who put a lot of pressure on teachers to "change our teaching to fill in gaps where scores have been low in the past." She gave the example of running records, where teachers test students' fluency and comprehension by doing a cold read. She heard that some teachers felt enough pressure by administration to demonstrate growth that they would show students the story in advance in order to improve their scores, though that is something she never felt the need to do personally. Naomi shared that she had been directed by an administrator to redo a student's post-assessment over and over until that student performed better, eventually boosting both of their overall evaluations. Some of the participants were not aware of the administrators' evaluation system at all, but upon learning about this in the inquiry group, several teachers had realizations. Nicole added "But, now knowing what principals are being evaluated on, that totally makes sense." By "that," Nicole was referring to the tendency of some administrators to ask students to redo assessments until the desired outcome was achieved. Nicole was suggesting that as the teachers in the study uncovered more information about how administrators were personally affected by teacher evaluation, the administrators' behaviors during the teacher evaluation process made more sense. To the

teachers in the study, the administrators were exerting pressure to demonstrate increase in student performance even if through questionable means.

However, the teachers in the study did not want to jump to conclusions on behalf of the administration. Brittany asked “Will you be getting perspectives of administrators? I wonder what they would want.” Dorothy added, “Are they happy with the way it is?” Here Dorothy was referring to the teacher evaluation system which teachers in the study frequently conflated with formal observation in our discussions. Brittany shared that one of her supervisors visits the classrooms every day and would probably rather skip the formal observation if he could and “only observe when he feels there is a need to.” Even though teachers felt that administrator bias had directly harmed many of them and made the system unfair, they still showed compassion for administrators. They recognized supervisors may also be caught in the system, and face constraints or factors out of their own control. They suspected that administration might have critiques of the teacher evaluation policy as well.

Just as teachers wondered whether administrators felt pressure to adhere to a curricular program, teachers in the study pointed out that administrators also seem constrained by the evaluation system, because it is based on teachers under their watch performing well. Nicole said “Principal evaluation or administrator evaluation have a piece where teacher evaluations go into that. So, there’s pressure on them to give us certain evaluations so that they can receive an evaluation. And it’s just...trickle down.” Some teachers went further than just administrators’ ratings and also held the perception that previous superintendent’s bonuses were dependent on teacher ratings. It turns out that this was a misconception of administrators in the district currently. After the state place a cap on superintendent salaries across New Jersey, this district offered opportunities for superintendents to earn additional stipends based on different measures.

Student growth goals were tied to the previous superintendents' opportunities for additional pay, but that practice did not continue into the present superintendent's contract.

### ***Unhelpful Feedback***

Another major finding is that teachers in the study generally did not receive feedback that improved student learning in their classroom and provided suggestions for how to make it better. Teachers were able to articulate specific critiques of the policy mandates based on their interaction with administrators at the local level during the formal observation cycle. They critiqued the level of feedback they received, felt evaluations were very affected by administrator bias and wanted more helpful feedback, but they did strive to understand the constraints administrators were under as well.

I set out to analyze the data to explore the ways in which the formal observation actually is instructive to the teachers in the study and found that it produced shallow outcomes—few to none associated with improved student learning. Therefore, it was implicitly instructing them to design lessons that matched the supervisors' preferences rather than helping them improve as teachers. For that reason, I assert that the hidden curriculum in the teachers' negative experiences with formal observation was that they should seek out their supervisors' preferences and put on, what Brittany called, “a great show” to receive high marks.

Teachers felt like the feedback they received from supervisors over the years was limited in its usefulness. Dorothy said, “I thought it was not really helpful in my opinion...I just noticed from the same supervisor the same response each time.” Mabel said if an administrator offered suggestions for improvement whether or not they genuinely felt the teacher needed it. She explained that it came across as lacking meaning, “it seemed like the supervisors felt that they had to give some area for improvement. I would've appreciated a different one...I wonder if they

have to give an area for improvement and they kind of fish for things artificially.” Several teachers, who had been in the profession for a while, gave examples of administrators hyper-focusing on classroom management in their feedback instead of an issue that was more meaningful to them as experienced teachers. In the vignette presented above, Naomi was grateful that the supervisor caught students misbehaving so she could address those particular issues later; however, that feedback was not helpful in deeper ways to improve her teaching. The supervisor simply acted as extra surveillance of the students’ behavior in that particular moment.

In her interview, Brittany discussed that she could see how teachers wonder where their marks come from. “It’s frustrating because they are like, ‘You didn’t tell me I have to change anything or improve, but you’re telling me that I’m not highly effective, so what is that about?’” She said the feedback has to be meaningful for the teacher if you want them to improve and the rubric does not provide enough details.

Not only were ratings influenced by administrators’ preferences, but so was their feedback. Rose said, in observations, it “depends on what they are looking for.” She spoke of a time where two administrators observed her and provided different feedback. Yet, interrater reliability training is required in the NJ law as training to minimize subjectivity across administrators. NJ law requires supervisors to participate in two co-observations, where they each evaluate the same teacher and must align on a singular summary of feedback. Rose was not aware of this requirement; instead, she assumed that the administrators were possibly evaluating each other too. Despite what was supposed to be an exercise in coming together to rate a teacher, the co-observation revealed more evidence to Rose that administrators vary in what they would like to see in a classroom. She found that the supervisors’ feedback exposed they had different foci depending on what was important to them.



## **Support, Optimism and Dreams**

### ***Teachers' Suggestions to Improve Feedback***

The major suggestion from teachers in the study for how to improve the teacher evaluation system was simply that administrators should offer more meaningful feedback to teachers in the process. Some teachers in the study discussed a specific supervisor who provided quick, affirming feedback and resources, and some teachers were optimistic about what the system could be, while others remained skeptical.

In the whole study, there was only one instance of teachers sharing positive experiences with feedback. Several teachers in the study shared stories about a supervisor who provided good feedback because she was quick to offer it, affirmed their existing practice without criticism, and provided resources later. They liked this supervisor's approach and wished that more supervisors offered helpful feedback in this way. Naomi shared, "I found what was really helpful is that when she would observe you, you would get feedback immediately. It was not just what your lesson was, she would give you suggestions and one of the suggestions she gave me was a worksheet about writing 'long' off a post-it. She called it writing a four post-it, an idea I could use the very next day. It helped me build upon the lesson the very next day. And I thought it was helpful." Dorothy agreed that was one of the supervisor's strengths. For Dorothy, the supervisor gave her helpful links and suggestions that were not a criticism. They were "resources going in a positive direction and how you as a teacher could build upon what you were doing." The teachers valued this targeted feedback to help them improve their instruction practice, but feedback like this was not the norm.

Teachers were optimistic and wanted to see that teacher evaluation could be a positive influence in a district. Dorothy stated, "Ideally, it should be to improve your effectiveness as a teacher...it would improve instruction overall for your students and on the grade level or in the

school. So, it should be a more positive sort of thing.” Just as teachers sought out the administrators’ perspectives compassionately, there were other moments in the study when teachers were clearly rooting for these systems to work. However, it was not clear whether adding on meaningful feedback would prevent the negative feelings teacher felt before and during the teacher evaluation process, or whether it would address the concerns about administrative bias. Though, their comments suggested that receiving meaningful feedback may make the experience of teacher evaluation feel like it was more worth their time. Brittany saw the potential for growth and suggested taking administrators out of the equation. She said, “I’d rather teachers observe other teachers and provide feedback informally. That’d be more beneficial.” To Brittany, since teachers did not have authority to cite or punish other teachers, their observations could truly be based on meaningful feedback to help each other grow. Brittany also recognized that there was potential for the administrator feedback to be better received if teachers changed their mindset. She said the teachers felt under attack, “every teacher feels like they’re coming in to catch you and that’s a problem. That mindset has to change.” Kimberly also suggested that teachers observe each other in action. She had observed another teacher and loved his calming environment and the way the students responded to him. It inspired her to do things differently in her own classroom. Teresa pointed out that formal observation does not help struggling teachers improve. Teresa said “I think some teachers do need it, but I think they need even more than what they are getting now. They need pop-in’s every two weeks.” Teresa felt the formal observations were inadequate ways to support struggling teachers who would benefit from regular support every two weeks. She was equating the formal observation with administrator contact, since some of the only contact teachers in the district have with administrators in their classroom is during that formal observation.

When asked about the purpose of the teacher evaluation system in follow-up interviews, teachers struggled to answer. Teresa said “That’s a good question. I don’t know.” Kimberly said her job would not be different if teacher evaluation were eliminated. She said, “I still want a smoothly run class, so I am still going to work on my classroom management.” Brittany said it would not change her teaching if it were eliminated but dreamed about an alternative where teachers participate in what she called “collaborative feedback” with each other instead. She described it as “not a judgement call. It’s a kind of work together and reflect.”

As mentioned, the teachers in the study were not inspired to take action as a result of our work together. However, they posed questions about the system that could inform further lines of inquiry and investigation should the group take up a research project in the future.

I have compiled their questions below and ordered them based on chronological appearance in the study:

- How are the state numbers calculated?
- Why do we do this?
- Why do we need to be evaluated at least using the parameters that they are doing and are they effective?
- Do you feel teacher evaluation has helped you grow as a teacher over the years?
- Do you think evaluations are fair?
- How do teachers feel?
- What do teachers feel their needs are?
- What can we do to better train new teachers coming into the profession and better support teachers who are already in the profession?
- Are administrators happy with the teacher evaluation policy?

## VI – DISCUSSION AND IMPLICATIONS

The purpose of this study was to utilize teacher research to understand and explain teacher evaluation policy mandates and their impact on teachers. I also documented qualitative methods informed by the values of teacher research to engage teachers as policy actors. As a practicing public elementary school teacher, I sought to broaden what Somekh and Zeichner (2009) allude to as a narrow view of teachers' work and demonstrate how teacher research can explore the political aspects of teachers' experience. In this study's final chapter, I offer a discussion of the study as well as implications for developing teachers as policy actors in teacher advocacy, professional development, and future research.

This exploration of teachers as policy actors was guided by two research questions: 1) What are the underlying discourses in the official documents of teacher evaluation policy? 2) What responses surface when teachers study teacher evaluation policies and practices in their New Jersey school district? Through document analysis of official teacher evaluation policy texts, I explored underlying discourses that target teachers in this reform effort. By facilitating a teacher inquiry group engaged in discussion about official texts and arts-based, storytelling activities, I explored teachers' meaning making of the policy, their feelings and attitudes, their critique, and their suggestions for improvement.

I start out with an overview of each chapter. Then, I discuss my findings within the larger context of the existing literature presented in Chapter II. Next, I attend to the implications for developing teachers as policy actors in teacher advocacy, professional development, and future research. Finally, I end with some concluding thoughts about teachers as policy actors.

## **Discussion**

Chapter IV uncovers the political discourses in official documents. By studying official documents of teacher evaluation policy, I was able to construct a timeline and introduction to the policy mandates, so that the teachers in my study could fully engage in the policy discourse. I explored an apparent theory of action based on money and global competitiveness manifest in these mandates, as well as a deeper intent to target teachers' tenure. Federal mandates operated under the premise that improving teacher quality closes the achievement gap, but, in reality, the mandates act as a measure to punish teachers. Packaged in the façade of a democratic decision-making process, they also purport to be crafted with and receptive to educator input. Through my document analysis, I found that these official documents are inaccessible to teachers, unresponsive to public comment, and contain standards that are federally, not locally, defined.

Chapter V demonstrates how teachers in the study were aware of underlying discourses based on their lived experience with teacher evaluation policy mandates in their school district. While they were rooting for teacher evaluation systems to help them improve their practice, they recognized that teacher evaluation was being used to monitor teachers with the threat of elimination. Teachers' recollections of episodes with formal observation inspired negative feelings. They felt stress around the formal observation cycle, and they characterized the observations as unrealistic and unfair. Teachers in the study also believed that other components of the teacher evaluation system, specifically those involving state learning standards and state tests, contained an oversimplified view of children that was incompatible with their own held beliefs. The teacher-administrator dynamic, under teacher evaluation, was affected negatively as well. Despite feeling removed from the policymaking process, teachers in the study supported each other, remained optimistic and dreamt up alternatives to ameliorate teacher evaluation.

Findings communicated in Chapter IV and V illustrate the negative effects of market-based reforms, like teacher evaluation policy mandates, on teachers in their New Jersey school district. My findings were in line with the existing literature on teachers' responses to reform. Such research has shown that reform can result in anxiety and dread (Evans, 2000; Troman 2000), and these effects were noted and discussed among participants in this study, as seven out of eleven participants reported feeling anxiety due to their job. The formal observation cycle and the lack of resources for students caused anxiety in teachers like Naomi and Melissa. Participants, like Brittany and Nicole, felt lack of trust for teachers, described as common in policy by Van den Berg and Schultz (2014) in their experiences with administration during teacher evaluation. They were likely to feel more trusted by other teachers and supervisors who were recently teachers. This study contributes to the research on teachers' responses to reforms.

Additionally, findings suggest a future direction in regards to the longstanding tension among different interpretations of the conceptions of teaching and administrator's roles in teacher evaluation. As described in Chapter 2, Darling-Hammond (1984) offers different administrative roles based on different conceptions of teachers, including teaching as labor, teaching as craft, teaching as profession, and teaching as art. The role of the administrator in teacher evaluation varies depending on the dominant conception of teaching at hand. Current approaches to teacher evaluation demonstrate the former conceptions, resulting in administrators who utilize the teacher evaluation system to monitor teachers. Moving to a conception of teaching as art (Gage, 1978) values personal insight with self-evaluation focused on holistic qualities could shift the administrator's role to someone who encourages teachers rather than monitoring or evaluating them. Contemporary policy suggests this conception works best to inform teacher evaluation for experienced teachers. However, due to the negative effects of the

current teacher evaluation system, which really favors the conception of teaching as labor, a broader use of this conception is warranted to address some of the harm found in that narrow conception of teaching.

This dissertation study's original contribution is that it offers the "autobiographical analysis," directly from and by teachers in their New Jersey school district (Bullough & Pinnegar, 2001 as cited in Cochran-Smith & Donnell, 2006). The other studies mentioned were designed by researchers who were not also full-time teachers. With the exception of Evans (2000) who took on part-time teaching duties in one of her three sample schools, the researchers worked as outsiders with teachers from different schools, and they utilized interviews or reflective assignments to gather their data. However, as a full-time elementary school teacher for over fifteen years in the school district where the study was conducted, I acted as a teacher-researcher. I designed this study and facilitated a teacher inquiry group among fellow teachers in the district to explore a policy issue that affects our daily work. Teacher-researchers are in a special position as insiders in the professional context to illuminate the distinction between "what is intended and what occurs," in practice (Cochran-Smith and Donnell, 2006). As a teacher-research project, this study offers a source of validity for teachers' professional practice in the face of teacher evaluation policy mandates. It also has transformative potential as the broader agenda of Cochran-Smith and Lytle's (1993) "teacher research" moves beyond the classroom and focuses on "educators' sense of social responsibility and social action in the service of a democratic society" (p.58).

As I reflected on my role as teacher-researcher, I became aware of the need to let go of my desire to be just another teacher among them in the study. This had been my original intention as I designed the study, as I truly wanted to see myself as on an equal footing with the

other participants. However, I had to acknowledge the power and agency I assumed in my additional roles i.e., research designer, lead researcher, “university partner,” (Fregeau and Leier, 2002), data analyst, and more. I entered the project with the intention that this would be a collaborative teacher research project, but the participants did not envision themselves in the role of co-researchers. They saw themselves as participating in *my* study. They found the collaborative discussions meaningful, were able to earn district credits toward release from instructional duties, and were also supporting a colleague. In their interviews, teachers reported that their favorite part of the experience was coming together in community. Making space for teacher inquiry groups like this one might support Little’s (1990) finding that teachers aiding one another and sharing regularly contribute to stability and job satisfaction. Although I wanted to engage in a teacher research group as an equal participant, it is evident that without my initiating the research project, the teacher inquiry group would not have existed.

My hope was also that the teachers would feel activated to mobilize and voice their concerns and recommendations for teacher evaluation, but this movement to collective action was not realized. Instead, teachers like Kimberly, Teresa and Dorothy, felt they did not have much power in changing things. They would consider joining onto a collective action, but it would need to already be in motion, have support from peers and possibly administrators too, and have clearly articulated goals for advocacy. They felt it would be challenging to gain consensus among such a large group, and were skeptical whether such action would be effective. They reacted supportively to the three case studies where teachers created change through teacher research, including Rust and Meyers’ (2006) use of TLNI to influence various local policy outcomes for teachers, Atkins’ (1988) study of teachers saving a teacher mentor program in California, and Fregeau and Leier’s (2002) discussion of using TERRA to address inequities.



However, their positive responses to the case studies and their participation in discussion around teacher evaluation in this study were not enough to move them to action.

If the collaborative teacher research had crossed over into the territory of action research, I would have utilized a critical lens. Koro-Ljungberg (2013) suggests “The movement between constructivist, social constructivist, and critical approaches can enable researchers to shift the epistemological focus from individual experiences to social meaning making and, ultimately, communal actions” (p. 772). This shifting worldview provides deeper understandings of how individual experiences are connected to societal structures. The connection between the individual’s experience and the societal structures would have been explored in more depth if the action research component had emerged. The prospect of a collective action around teacher evaluation was exciting to me as a teacher activist (Sachs 2000), but it was not realized.

### **Implications for the Study**

The findings from this study could be useful in several contexts. First, professional development could incorporate teacher inquiry groups as reflective collaborative tools to build community and to make space for teacher input on a district’s policies and procedures. Professional development could also provide instruction for teachers and administrators regarding navigation of said policy. Second, this study stresses the need for including lived experiences and narrative methods. Future research on teachers studying policy needs to include their lived experiences. in collecting data with teachers studying policy. Finally, teacher advocacy work could benefit from this study because it suggests we can develop teachers as activists, examine unintended consequences of policy, and stress the need for policy to depend on practice.

## **Implications for Professional Development**

Professional development strategy in school districts could incorporate teacher inquiry groups as reflective collaborative tools. Having time for teachers to come together in inquiry was not only satisfying for the teachers in the study, but they also reported it as the most memorable part of the work. Being in community was more meaningful than the topic of their discussion. This echoes the finds of Espiritu, Meier, Villazana-Price, & Wong (2002) who found that teachers coming together in research groups about language and literacy in early childhood helped them “move beyond the traditional physical and professional boundaries of [their] work” and combated the isolation they felt. The well-known sociologist Lortie (1975) wrote about the “cellular,” (p. 23) nature of teaching many years ago, where the lone teacher is often autonomous in her classroom, and little has changed in the time since. Social relationships have become separated from their context, with less likelihood to support basic trust, which requires openness and warmth between two face-to-face individuals (Troman, 2000). Teachers’ health and well-being are suffering the consequences of this lack of connection with other professionals, with a rise of anxiety and dread during high-risk times of reform. Professional development that incorporates the use of teacher inquiry groups to come together throughout the school can combat some of these ill effects of reform mandates.

Additionally, the participants felt they didn’t have time to think critically about issues. Incorporating collaborative teacher inquiry groups into professional development would build in time for what Dewey calls the “freed activity of mind,” without the need for so-called outside experts (Lagemann, 2002, p. 50). As Tracy pointed out during the study, administrators feared giving the teachers in the district “free” time, but, in reality, she said, “it is one of our most productive times.” Brittany also suggested that the district could build in reflection time each day where teachers could talk to each other or coaches rather than work on the endless list of

administrative assignments she called “extra baloney.” Stenhouse (1975) recognized that teachers have a commitment to self-questioning and improvement in their practice, and Brittany believed the majority of teachers in the district were “highly motivated, highly reflective and just want to get better always.”

Just as teachers in this study generated several suggestions that would improve teacher evaluation policy, teachers working together in another setting may provide helpful suggestions for a district’s policies and practices. In this study, teachers expressed the sentiment that feedback during the formal observation cycle could be more helpful. Naomi recalled a time when a supervisor provided quick feedback and resources, she could put to work in her classroom right away, and other teachers, like Dorothy, agreed that this supervisor was skilled in this way. Brittany valued feedback from administrators who took time to get to know her outside of the observation. Building that relationship meant a great deal to her. Additionally, several teachers, like Teresa, suggested reducing the frequency of stress-inducing, formal observations to one time per year for nontenured teachers. Teachers have specialized knowledge grounded in their daily work and are operating within an educational context that frames their experience. Valuing teachers’ knowledge and making space to engage them in collaborative inquiry group discussions as part of professional development initiatives, has the potential to generate helpful suggestions for school districts’ policies and procedures. Thus, the results of this study suggest that professional development efforts should include time for teachers to come together and collaborate in inquiry group discussions to benefit the teachers and improve the district.

Professional development could also include instruction about what policy is and how to navigate it as a practitioner for both teachers and administrators. As discussed in Chapter 4, the inaccessibility of official policy documents prevents teachers from participating more fully.

Inquiry study groups help teachers cross what Star (2010) calls a “boundary” held by the official text of policy documents in order to improve conditions. Results from this study showed that teachers were not familiar with the concept of policy in general, and even after the researcher presented a lecture about teacher evaluation policy, they were still unclear. Teachers conflated the teacher evaluation policy with only one aspect of it, the formal observation, though they had criticisms of the use of student growth measures. They spoke about the use of state learning standards and standardized tests as if they were separate from teacher evaluation policy when they are a component of the teacher’s final rating. Teachers were also unsure who policymakers were other than elected officials and how they could influence policy. They discussed and dismissed the power of attending Board of Education meetings and expressed the belief that parents had more impact there. Teresa and Brittany suggested teachers and administrators would have to go “higher” to make an impact. Eventually, the group agreed with Rose that voting in elections was the best route to impact policy. Looking back to the underlying discourses in federal teacher evaluation policy, there are numerous market forces influencing policy decisions through lobbying. Teachers did not feel connected to lobbying as a way to make change in policy. Educating teachers and administrators on ways to get involved in lobbying efforts through professional development might help push their voices towards the center of policy decision-making. This refers to sustained and unified efforts by teachers and administrators across the school district, state and country. Policy is more likely to influence practice if it “cultivates practitioners’ use of the policy” (Cohen, Moffit, and Goldin, 2007). This requires input from educators as the key decision-makers to which policymakers should turn. If districts equip educators to understand and get involved in policy, the initiatives are more likely to be used with fidelity, rather than if resources are just allocated to a school system as a transaction.

## **Implications for Teacher Research and Policy**

Researchers may automatically think that teacher research is only about inquiries confined to the classroom. This qualitative study aimed to expand the typical conception of teacher research beyond classroom walls by amplifying teachers' knowledge in a broader policy context. As "systematic, intentional inquiry by teachers," (Cochran-Smith and Lytle, 1993, p.5), teacher research can be useful beyond the local context of a teacher's classroom, such as to inform research about policy. Cochran-Smith and Lytle (1993) argue that "teachers' roles as theorizers, interpreters, and critics of their own practice" have been ignored and "they have rarely been included in decisions about research as knowledge generation" (p. 1).

Given that teachers are considered central to reform, teachers initiating research projects about policy implementation, such as in this study, provides a meaningful way to analyze policy. One strategy that Cheung et al. (2010) suggests is considering the policy's alignment, or misalignment, between its aims and outcomes. Teachers in this study, like Rose, Dorothy and Nicole utilized a similar approach as they considered teacher evaluation policy mandates. They identified that the teacher evaluation policy had a stated purpose of helping teachers improve their practice after learning about the verbiage used in official policy documents. The stories they shared presented data that showed the mandate was not meeting its stated goals in their lived experience. Instead, in their eyes, the mandate acted like "the school district's insurance policy" to monitor, cite, and possibly eliminate teachers. Rose also discussed how teacher evaluation was presenting as more and more punitive over the years. It was not the fear of losing one's job that inspired negative feelings, but it was the lack of value and trust inherent in the ways in which the teacher evaluation system was working.

Findings like these are grounded in teachers' authentic perspectives. Only teachers have the unique insider perspective as the target of reforms, like teacher evaluation, so to fully capture their unique perspectives, they must be in control of the research from start to finish. In a time when teachers feel anxiety and distrust, having teachers in control of the research from start to finish, i.e., design the research plan, facilitate it, and interpret its findings, has the potential to offer rich data and unique insider insights that may not be gathered from an outside researcher. Validity in teacher research is established through "significance, quality, grounding, and the authority of autobiographical analysis" (Bullough & Pinnegar, 2001 as cited in Cochran-Smith & Donnell, 2006). As analysis of data is primarily orally constructed and storytelling is a major feature of teacher research, we must continue to advocate for the valuing of teachers' knowledge and oral methods of analysis in the research community.

As argued in this section, teacher research is a potential methodological choice that values and captures teachers' specialized knowledge to contribute to the broader policy discussion.

### **Implications for Teacher Advocacy Work**

According to Hirsch (2016), one signal of the end of public education is policymaking that is based on unelected, private individuals who are unaccountable to the public. He argues that educational policy is now constructed outside of federal and state governments, under the influence of many outside groups who hold market fundamentalist ideology. Hirsch argues that corporate reform efforts, like teacher evaluation policy mandates, are a way to address inequality in society without confronting the societal inequality caused by neoliberal capitalism. He also suggests that reforms are constantly justified based on a "manufactured crisis" (p.108) instead of the real crisis caused by neoliberal capitalism. Given that the policymaking process is determined

by outside groups, teachers must rise up together among them or in opposition to them, if needed. Teachers need to work in coalition with other stakeholders, through advocacy efforts, to make policymakers truly care. Teachers should utilize existing organizations or create new ones that advocate for their interests, in order to preserve their profession and reclaim democratic decision-making in public education.

In this study, teachers did not express a desire to organize themselves into collective action based on their critical thoughts about the issue of teacher evaluation. Recruiting participants was relatively successful; I worked within the schedule of the school district and did not require teachers to give up any additional time. My familiarity with the participants also helped earn their trust. While my understanding of the work firsthand as a teacher beside them was key to organizing and directing the group to interrogate seemingly natural structures within our district. These claims are supported by the data which showed authentic responses with real critique from teachers.

The short duration of time that we met during the teacher inquiry groups limited our ability to organize action. Meeting in a workshop format for three days within a short period of time allowed us to learn about teacher policy together and discuss important critiques. However, it was not enough to activate the participants. Additionally, teachers in the group, like Teresa and Naomi, said in interviews that if many people came together on a unified front alongside administrators, then they would join in the action to address teacher evaluation concerns. Teresa gave the example of a successful organized effort by her teacher's union to attend and speak out at Board of Education meetings regularly to settle their disputed contract. Brittany pointed out that successful action would have to be "more of a collaborative effort to come up with a solution rather than a blame game." Therefore, having more participants in the group or building

coalitions with administrators, or leveraging the already existing teachers' union could be avenues to make the teachers feel more powerful as part of a bigger group.

In these sessions, teachers shared negative responses to teacher evaluation and many critiques; but their strong reactions were not enough to ignite the participants into taking action. Teachers in the study felt stress or worry around the formal observation cycle. They critiqued the level of feedback they received, administrators' subjectivity, and the ways students were oversimplified by teacher evaluation mandates. Despite all of these critiques, they did not desire to take action to try to change the teacher evaluation system. As Dorothy pointed out in her interview, if she did not like something strongly, she would rather leave the job than work to change the conditions. And, since this study was conducted, I have observed increasing numbers of teachers resigning from the district, leaving the profession or retiring, in some cases, early. However, not everyone has the privilege of being able to leave their job. While the study's findings provide guidance around what teachers might need to feel supported in an organized effort, I am left with the question of what would ignite teachers to take action at the grassroots level.

There are several directions teacher advocacy could pursue as a result of these findings. Teachers could offer suggestions, listed above, that would improve their school district's policies, with implications for state and federal changes. Additionally, teacher advocates could advocate to reduce the harm created by the teacher evaluation system. As the object of the evaluation reform mandates, teachers are in an important position to address this. Teachers in the study suggested reducing the number of observations conducted per year, allowing teachers to veto an observation if it did not go well, and having teachers offer feedback to each other instead of administrator feedback. Advocacy efforts fueled by teachers' solutions like these, could help



improve these systems to reduce the harm that has been created until more liberatory practices can be created. While a reimagined future may include the elimination of teacher evaluation, this may not be achievable overnight, and the harm that teacher evaluation mandates cause must be addressed in the meantime.

### **Critique of the Study**

During my dissertation defense, the question arose as to whether I influenced the participants into critiquing the teacher evaluation. Aside from teachers sharing positive feedback they received from a favored supervisor, the majority of the recollections and responses to teacher evaluation were negative. After revisiting the data, I could not conclude any significant

As a teacher activist, I also had hopes for the potential of this project to be an authentic teacher research project that would evolve into action research. I designed it as a collaborative teacher research project, but I did not design it *with* the participants. In the end, I was alone driving the structure of each session, the questions, the focus of the project, and the analysis of the data. So, perhaps it is more fitting to refer to the project methodology as a qualitative study informed by the values of teacher research rather than a true teacher research project. I was a teacher leading research about teacher evaluation policy mandates, which are deeply ingrained in teachers' work, but the participants and I were not co-researchers.

### **Concluding Remarks**

My experiences as a researcher in this study has been illuminating for me. I have spent almost twenty years as an elementary school teacher. I spend most of my time with children. To be able to explore another aspect of teachers' work with in such depth alongside other teachers has been personally satisfying. In this final section, I would like to share my learnings from this work. I have learned that the official governmental policy documentation is truly a genre in itself, and it takes a lot of effort to navigate. My preparation as a teacher did not include much

discussion of this genre; however, these documents affect teachers greatly. They have lasting effects on our daily lives as practitioners, so it is worthwhile to be able to access them. I utilized my experience as a doctoral student, familiarity with electoral politics and my activist background to navigate these documents. I recognize that I also had a desire to read them and unpack them, but assumed that many teachers do not have that interest. Through my work with teachers in this study, I learned to challenge that assumption. I was moved alongside the teachers in the study to build an understanding of teacher evaluation policy, unpack the language and contrast it with our lived experiences. So many studies are conducted by outsiders investigating the effects of policy mandates. Equipping teachers who are interested in leading their own study about the implementation of some policy offers the potential for more truthful and meaningful data. Additionally, teachers must rise up together to advocate for themselves.

As I move forward with the experience of this dissertation work, I want to use my experiences to continue to advocate for teachers. I have learned how to navigate official policy documents, and how to plan, execute, and analysis data from a study. I hope to utilize these tools for the betterment of teachers and students in the field of education.

## References

- 6, Perri (2012). When forethought and outturn part. In H. Margetts, Perri 6., & C. Hood, *Paradoxes of modernization: Unintended consequences of public policy reform*. Oxford, England: Oxford University Press.
- Anderson, L.M., Butler, A., Palmiter, A. & Arcaira, E. (2016). *Study of emergent teacher evaluation systems: A pilot study*. United States Department of Education.  
<https://www2.ed.gov/rschstat/eval/teaching/emerging-teacher-evaluation/report.pdf>
- Ashendon, D., Connell, B., Dowsett, G., & Kessler, S. (1987). Teachers and working-class schooling. In Livingston, D.W., *Critical pedagogy and cultural power*. Palgrave, London.
- Atkin, J.M. (1994). Teacher research to change policy. In Hollingsworth, S. and Sockett, H., *Teacher research and educational reform: 93rd yearbook of the National Society for the Study of Education*. Chicago: University of Chicago Press.
- Apple, M.W. (2001). Comparing neo-liberal projects and inequality in education. *Comparative Education*, 37(4), 409-423.
- Ball, S.J. Maguire, M., Braun, A., & Hoskins, K. (2011). Policy actors: doing policy work in Schools. *Discourse: Studies in the Cultural Politics of Education*, 32(4), 625-639.
- Barber, B. R. (1984). *Strong democracy: Participatory politics for a new age*. Berkeley, CA: University of California Press.
- Bell, L. (2010). The storytelling project model: A theoretical framework for critical examination of racism through the arts. *Teachers College Record*, 112(9), 2295-2319.
- Berliner, David C. (2009). *Poverty and Potential: Out-of-School Factors and School*

- Success*. Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. <http://epicpolicy.org/publication/poverty-and-potential>
- Bernard, H.R. (2011). *Research methods in anthropology: Qualitative approaches*. Blue Ridge Summit: AltaMira Press.
- Bogdan, R.C. & Biklen, S.K. (2016). *Qualitative research for education: An introduction to and methods*. Boston, Massachusetts: Pearson.
- Boyd, D., Grossman, P., Ing, M. Lankford, H., Loeb, S., & Wyckoff, J. (2011) The influence of school administrators on teacher retention decisions. *American Educational Research Journal*, 48(2), 303-333.
- Bredeson, P. V., & Kose, B. W. (2007). Responding to the education reform agenda: A study of school superintendents' instructional leadership. *Education Policy Analysis Archives*, 15, 5-5.
- brown, a. m. (2017). *Emergent Strategy*. Edinburgh, Scotland: AK Press.
- Calefati, J. (2011, December 23). N.J. receives \$38 million in Race to the Top education funding. *The Star Ledger*.  
[https://www.nj.com/news/2011/12/nj\\_receives\\_38\\_million\\_in\\_race.html](https://www.nj.com/news/2011/12/nj_receives_38_million_in_race.html)
- Carbno, W. C. (1981). Student evaluation of teacher effectiveness: A case study. *Educational and Psychological Measurement*, 41.
- Cheung, K.K., Mirzaei, M., & Leeder, S. (2010). Health policy analysis: A tool to evaluate in policy documents the alignment between policy statements and intended outcomes. *Australian Health Review*, 34(4), 405-413. doi: 10.1071/AH09767

- Cochran-Smith, M. & Fries, M.K. (2001). Sticks, stones, and ideology: The discourse of reform in teacher education. *Educational Researcher*, 30(8), 3-15.
- Cochran-Smith, M., & Lytle, S. L. (1993). *Inside: Teacher research and knowledge*. New York, NY: Teachers College Press.
- Cochran-Smith, M. and Donnell, K. (2006). Practitioner inquiry: Blurring the boundaries of research and practice. In Green, J. L., Camilli, G., Elmore, P. B., & American Educational Research Association. *Handbook of complementary methods in education research*. Mahwah, N.J: Lawrence Erlbaum Associates.
- Cochran-Smith, M., Stern, R., Sánchez, J.G., Miller, A., Keefe, E.S., Fernández, M.B., Chang, W., Carney, M.C., Burton, S., & Baker, M. (2016). *Holding teacher preparation accountable: A review of claims and evidence*. Boulder: National Education Policy Center. <http://nepc.colorado.edu/publication/teacher-prep>
- Cohen, D. K., Moffitt, S. L., & Goldin, S. (2007). Policy and practice: The dilemma. *American Journal of Education*, 113(4), 515-548.
- Corcoran, S. P. (2010). *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures of teacher effectiveness in policy and practice*. Providence, RI: Annenberg Institute for School Reform, Brown University.
- Cuban, L., & Tyack, D. B. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.
- Cuban, L. (2010, April 3). Parsing the meaning of school reforms [Blog post]. <https://larrycuban.wordpress.com/2010/04/03/parsing-the-meaning-of-school-reforms/>

- Danielson, C. (2008). *The handbook for enhancing professional practice: Using the framework for teaching in your school*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Darling-Hammond, L. (2012). Creating a comprehensive system for evaluating and supporting effective teaching. Stanford Center for Opportunity Policy in Education (SCOPE).  
<https://edpolicy.stanford.edu/sites/default/files/publications/creating-comprehensive-system-evaluating-and-supporting-effective-teaching.pdf>
- Durkheim, E. (1956). *Education and sociology*. Glencoe: Free Press.
- Eglash, Bennett, & O'Donnell. (2006). Culturally situated design tools: Ethnocomputing from field site to classroom. *American Anthropologist*, 108(2), 347-362.
- Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, 3(4), 99-117.
- Espiritu, E., Meier, D.R., Villazana-Price, N., & Wong, M.K. (2002). A Collaborative Project on Language and Literacy Learning: Promoting Teacher Research in Early Childhood Education. *Young Children*, 57(5), 71-78.
- Evans, L. (2000). The effects of educational change on morale, job satisfaction and motivation. *Journal of Educational Change*, 1(2), 173-192.
- Firestone, W.A. Nordin, T.L. , Shcherbakov, A., Kirova, D., and Blitz, C.L. (2014, January). *New Jersey Pilot Teacher Evaluation Program: Year 2 Final Report*. New Brunswick: Rutgers Graduate School of Education.  
<https://www.nj.gov/education/AchieveNJ/resources/FinalRUGSEReport.pdf>

NJ Exec. Order No. 175, (2020). <https://nj.gov/infobank/eo/056murphy/pdf/EO-175.pdf>

NJ Exec. Order No. 214, (2020). <https://nj.gov/infobank/eo/056murphy/pdf/EO-214.pdf>

Fals-Borda, O. (2001). Guest editorial: From Cartagena to Ballarat: A report on the Joint Fifth World Congress on action learning, action research, and process management and Ninth World Congress on participatory action research. *Systemic Practice and Action Research*, 14(2), 125.

Firestone, W.A., Blitz, C.L., Gitomer, D.H. Kirova, D., Shcherbakov, A., & Nordon, T.I. (2013). *New Jersey Teacher Evaluation RU-GSE External Assessment, Year 1 Report*. New Brunswick, NJ: Rutgers Graduate School of Education.  
[www.nj.gov/education/AchieveNJ/resources/RUGSE11-12.pdf](http://www.nj.gov/education/AchieveNJ/resources/RUGSE11-12.pdf)

Fregeau, L. A., & Leier, R. D. (2002). Praxis and teacher vision of socially just school reform. In J. J. Slater, S. M. Fain & C. A. Rossatto (Eds.), *The Freirean legacy: Educating for social justice*. New York: Peter Lang.

Freire, P. (1972). *Pedagogy of the oppressed*. London: Sheed & Ward.

Galman, S. C., & Mallozzi, C. A. (2012). She's not there: Women and gender as disappearing foci in US research on the elementary school teacher, 1995–Present. *Review of Educational Research*, 82(3), 243-295.

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ.

Gitomer, D.H. Martinez, J.F., Battey, D. & Hyland, N.E. (2019). Assessing the assessment: Evidence of reliability and validity in the edTPA. *American Educational Research*

*Journal.*

Given, L. M. (2008). *The SAGE encyclopedia of qualitative research methods*. Thousand Oaks, CA: SAGE Publications Ltd.

Goodland, J. (2008). Agenda for education in democracy. *Journal of Educational Controversy*, 3(1).

Goodson, I., Moore, S., & Hargreaves, A. (2006). Teacher nostalgia and the sustainability of reform: The generation and degeneration of teachers' missions, memory, and meaning. *Educational Administration Quarterly*, 42(1), 42-61.

Grant, A. (2019). *Doing excellent social research with documents: Practical examples and guidance for qualitative researchers*. New York, NY: Routledge.

Groundwater-Smith, S. & Sachs, J. (2002). The activist professional and the reinstatement of trust. *Cambridge Journal of Education*, 32(3).

Hall, G. E., & Hord, S. M. (1981). A national agenda for research in teacher education for the 1980s. *Journal of Teacher Education*, 32(2), 4-8.

Hanushek, E. A., & Rivkin, S. G. (2006). Teacher quality. In Hanushek, E.A. and Welch, F. (Eds.), *Handbook of the economics of education*. (pp. 1052-1078). Amsterdam, Netherlands: North Holland.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1), 84-117.

Hatch, J. A. (2002). *Doing qualitative research in education settings*. Suny Press.

Hargreaves, A. (1998). The emotional practice of teaching. *Teaching and Teacher Education*,



14(8), 835-54.

Harvard Graduate School of Education. (n.d.). A user's guide to peer assistance and review. The

Project on the Next Generation of Teachers. <http://www.gse.harvard.edu/~ngt/par/>

Heath, S. B., & Street, B. V. (2008). *On ethnography: Approaches to language and literacy research*. London, England: Routledge.

Hengst, F.E. (2018) Teachers as policy actors: A collaborative teacher research project exploring teacher evaluation in Flowertown (Unpublished doctoral dissertation proposal). Teachers College, New York, N.Y.

Hollingsworth, S. (1994). Introduction. In S. Hollingsworth, & H. Sockett (Eds.), *Teacher research and educational reform*. Chicago, IL: NSSE.

Hollingsworth, S., & Miller, J. L. (1994). Chapter VII: Rewriting "gender equity" in teacher research. *Teachers College Record*, 95(6), 121-140.

Hirsch, D.W. (2016). *The end of public schools: The corporate reform agenda to privatize education*. New York: Routledge.

Ingersoll, R. (2001). Teacher turnover and teacher shortages: An organizational analysis.

[http://repository.upenn.edu/gse\\_pubs/94](http://repository.upenn.edu/gse_pubs/94)

Jansen, J. D. (2001). Image-ining teachers: Policy images and teacher identity in South African classrooms. *South African Journal of Education*, 21(4), 242–246.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education review*, 27(6), 615-631.

- Koro-Ljunberg, M. (2013). A methodological framework for studying policy oriented teacher \ inquiry in qualitative research contexts. *Teachers and Teaching*, 20(6), 764-782.
- Lagemann, E. C. (2000). *An elusive science: The troubling history of education research*. Chicago, IL: University of Chicago Press.
- Lazovick, R. M. (2020). *A Grounded Theory of the Impact of TEACHNJ on the Instructional Leadership of Superintendents*. [Doctoral dissertation, Rowan University]. ProQuest Dissertations & Theses Global.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Lortie, D. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press.
- Love, B. (2019). *We want to do more than survive: abolitionist teaching and the pursuit of educational freedom*. Boston, Massachusetts: Beacon Press.
- Love, P. (2003). Document Analysis. In Stage and Manning, *Research in the College Context*. New York, NY: Brunner-Routledge.
- Lytle, S.L. & Cochran-Smith, M. (1992). Teacher research as a way of knowing. *Harvard Educational Review*, 62, 447-474.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. Sage publications.
- McTaggart, R. (1997). *Participatory action research: International contexts and consequences*. Suny Press.
- Milanowski, A. (2011). Strategic measures of teacher performance. *Phi Delta Kappan*, 92(7), 19-25.

May, R. (1994). *The courage to create*. New York: WW Norton & Co Milanowski, A. (2011).

Strategic measures of teacher performance. *Phi Delta Kappan*, 92(7), 19-25.

National Commission on Teaching and America's Future. (2007). *Policy brief: The High Cost of Teacher Turnover*. Washington, DC.

National Conference of State Legislators. (2016, August). *No time to lose: How to build a world-class education system state by state*.

[www.ncsl.org/documents/educ/Edu\\_International\\_Final\\_V2.pdf](http://www.ncsl.org/documents/educ/Edu_International_Final_V2.pdf)

National Council on Teacher Quality. (2013). *Connect the dots: Using evaluations of teacher effectiveness to inform policy and practice*.

[https://www.nctq.org/dmsView/State\\_of\\_the\\_States\\_2013\\_Using\\_Teacher\\_Evaluations](https://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations)

[NCTQ\\_Report](#)

National Council on Teacher Quality. (2019). *State of the states 2019: Teacher & principal evaluation policy*.

[https://www.nctq.org/dmsView/NCTQ\\_State\\_of\\_the\\_States\\_2019\\_print\\_ready](https://www.nctq.org/dmsView/NCTQ_State_of_the_States_2019_print_ready)

National Education Association (2012). *Transforming teaching: Connecting personal responsibility with student learning* [white paper].

<http://www.nea.org/assets/docs/Transformingteaching2012.pdf>

National Policy Board for Educational Administration (2015). Professional Standards for

Educational Leaders 2015. Reston, VA: Author. [https://www.npbea.org/wp-](https://www.npbea.org/wp-content/uploads/2017/06/Professional-Standards-for-Educational-Leaders_2015.pdf)

[content/uploads/2017/06/Professional-Standards-for-Educational-Leaders\\_2015.pdf](https://www.npbea.org/wp-content/uploads/2017/06/Professional-Standards-for-Educational-Leaders_2015.pdf)

National Center on Education and the Economy. (2007). *Tough choices or tough times: The*

*report of the new commission on the skills of the American workforce.* Jossey-Bass.

New Jersey Administrative Code (NJAC) Chapter 10 Law N.J.A.C. 6A:10, Educator

Effectiveness (2013) <https://www.state.nj.us/education/code/current/title6a/chap10.pdf>

New Jersey Department of Education (2013-2014). *Achieve New Jersey Final Implementation Report 2013-2014.*

<https://www.nj.gov/education/AchieveNJ/resources/201314AchieveNJImplementationReport.pdf>

New Jersey Department of Education (2015). *Excellent Educators for All: Draft pending U.S. Department of Education Review.*

<https://www2.ed.gov/programs/titleiparta/equitable/njequityplan060115.pdf>

New Jersey Department of Education (2015, May 1). *New Jersey Department of Education: Approved teacher practice evaluation instruments as of May 1, 2015.*

<http://www.nj.gov/education/AchieveNJ/teacher/approvedlist.pdf>

New Jersey Department of Education (2017, August). *Every Student Succeeds Act: New Jersey State Plan.* [www.nj.gov/education/ESSA/plan/plan.pdf](http://www.nj.gov/education/ESSA/plan/plan.pdf)

New Jersey Department of Education (2017, October). *Teacher evaluation and support.* Trenton, NJ. <http://www.state.nj.us/education/AchieveNJ/intro/1PagerTeachers.pdf>

New Jersey Department of Education. (2018, July 12) *New Jersey Department of Education Statewide Assessment Outreach: Summary of Findings, Recommendations for Next Steps.*

<https://www.nj.gov/education/assessment/docs/outreach/AssessmentReportSummary.pdf>

New Jersey Department of Education. (2020, March 16) Document 52 NJR 498(a) Proposed

Readoption with Amendments: N.J.A.C. 6A:10. *New Jersey Register*. 52 (6).

New Jersey Department of Education. (2021, October 4) Document 53 NJR 1676(a) Rule

Adoptions N.J.A.C. 6A:8-1.3, 3.1, 4.1, 4.3, 4.5, 5.1 and 5.2. *New Jersey Register*. 53

(19).

Nisbet, S., Dole, S., & Warren, E. (1997). Cooperative professional development for

mathematics teachers: A case study. In *People in Mathematics Education: Proceedings of the 20th Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 369-375). Mathematics Education Research Group of Australasia Inc.. –

double check this one

Noddings, N. (2007). *When school reform goes wrong*. New York, NY: Teachers College,

Columbia University Press.

Oakes, J., Franke, M. L., Quartz, K. H., & Rogers, J. (2002). Research for high-quality urban

teaching. *Journal of Teacher Education*, 53(3), 228-234.

Paine, D., Beal-Alvarez, J., & Scheetz, N. (2016). edTPA you have to attack this like a lion

hunting a gazelle on the Serengeti: Student perspectives on the implementation of a new initiative. *Issues in Teacher Education*, 25(2), 149-166.

Polikoff, M.S. (2015). The stability of observational and student survey measures of teaching

effectiveness. *American Journal of Education*, 121(2), 183-212.

Prior, L. (2011). *Using Documents in Social Research*. Los Angeles, CA: Sage.

- Rabinow, P., & Sullivan, W. M. (Eds.). (1987). *Interpretive social science: A second look*. University of California Press.
- Robertson-Kraft, C., & Zhang, R.S. (2016). Keeping great teachers: A case study on the impact and implementation of a pilot teacher evaluation system. *Educational Policy*,
- Rothstein, J., & Mathis, W. J. (2013). Review of “Have We Identified Effective Teachers?” and “A Composite Estimator of Effective Teaching”: Culminating findings from the Measures of Effective Teaching Project. Boulder: National Education Policy Center.
- Rust, F., & Meyers, E. (2006). The bright side: Teacher research in the context of educational reform and policy-making. *Teachers and Teaching*, 12(1), 69–86.
- Ryan, P. (2015). Positivism: paradigm or culture?, *Policy Studies*, 36(4), 417-433,
- Sachs, J. (2000). The activist professional. *Journal of Educational Change*. 1(1), 77-95.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. London, England: Temple Smith.
- Schubert, W., Marshall, J., Sears, J., Allen, L., & Roberts, P. (2007). *Turning points in curriculum: A contemporary American memoir*. Upper Saddle River, N.J.: Pearson/Merrill Prentice hall.
- Schuller, K. (2021). *The trouble with white women: A counterhistory of feminism*. Bold Type Books.
- Simons, H. and McCormick, B. (2007). Integrating arts-based inquiry into evaluation methodology: Opportunities and challenges. *Qualitative Inquiry*, 13(2), 292-311.
- Somekh, B. and Zeichner, K. (2009). Action research for educational reform: Remodelling action research theories and practices in local contexts. *Educational Action Research*,

17(1), 5-21.

Spring, J. (1989). *The Politics of American Education*. Addison-Wesley Longman Incorporated.

Stage, F.K. and Manning, K. (2003). What is Your Research Approach? In Stage, F.K. and Manning, K. *Research in the College Context: Approaches and Methods*. New York, NY: Routledge.

Stapleton, S. R. (2021). Teacher participatory action research (TPAR): A methodological framework for political teacher research. *Action Research*, 19(2), 161-178.

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology, & Human Values*, 35(5), 601–617. <http://www.jstor.org/stable/25746386>

Stenhouse, L. (1975). *An introduction to curriculum research and development*. London: Heinemann Educational Books.

Stenhouse, L. (1978). Case Study and Case Records: towards a contemporary history of education. *British Research Journal*, 4(2).

The TEACHNJ Act, NJ S1455/NJA3343, 215<sup>th</sup> Legislature. (2012).

[www.njleg.state.nj.us/2012/Bills/PL12/26\\_.pdf](http://www.njleg.state.nj.us/2012/Bills/PL12/26_.pdf)

The White House. (2012) “Race to the Top” Education: Knowledge and Skills for the Jobs of the Future. Archived at <https://obamawhitehouse.archives.gov/issues/education/k-12/race-to-the-top>

Troman, G. (2000). Teachers 'stress in the low trust society. *British Journal of Sociology of Education*, 21(3), 331-353.

United Federation of Teachers (2010, October 20). UFT to file suit to prevent release of incorrect

- teacher data. *UFT.org*. <http://www.uft.org/press-releases/uft-file-suit-prevent-release-incorrect-teacher-data>
- U.S. Department of Education. (2011-2012). *Educator Equity Profile: New Jersey*  
<https://www2.ed.gov/programs/titleiparta/equitable/njeep.pdf>
- U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2010).  
*ESEA Blueprint for Reform*. <https://www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf>
- van den Berg, G., & Schulze, S. (2014). Teachers' sense of self amid adaptation to educational reform. *Africa Education Review*, 11(1), 59-76.
- Vygotskiĭ, L. S., Rieber, R. W., & Carton, A. S. (1987). *The collected works of L. S. Vygotsky*.  
New York, NY: Plenum Press.
- Walker, M. (2007). Action research and narratives: 'Finely Aware and Richly Responsible'  
*Educational Action Research* 15(2), 295–303
- Weisberg, D., (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York, NY: TNTP.  
<https://tntp.org/publications/view/the-widget-effect-failure-to-act-on-differences-in-teacher-effectiveness>
- Xu, X., Grant, L. W., & Ward, T. (2016). Validation of a statewide teacher evaluation system: Relationship between scores and from evaluation and student progress. *NASSP Bulletin*, 100(4), 203-222.
- Zeichner, K. (1991). Contradictions and tensions in the professionalization of teaching and the democratization of schools. *Teachers College Record*, 92(3), 363-379.



## Appendix A: Follow-up Interview Protocol

**Protocol Title:** Teachers as Policy Actors: A Collaborative Teacher Research Project Exploring Teacher Evaluation in Flowertown

**Principal Investigator:** Filomena Hengst, Teachers College Doctoral Student

### Follow-up Interview Protocol<sup>11</sup>

Thank you for consenting to participate in the study called “Teachers as policy actors: A collaborative teacher research project exploring teacher evaluation in Flowertown”. This interview will be recorded, reviewed by myself and kept in a password protected storage service. After my transcription of the conversation is complete, the audio files will be destroyed, and transcriptions will be stored in a password protected storage service only accessible to myself. Your identity will be kept confidential as pseudonyms will be assigned in any publications arising out of this work; however a knowledgeable reader may be able to infer your identity from the context. I would also like to remind you that you have the option to withdraw at any time, if desired.

#### *Interview Guide:*

1. Since the topic of this interview is teacher evaluation, let me start by asking you what is your definition of teacher evaluation? Please describe some examples of teacher evaluation.
2. What was your earliest experience with teacher evaluation? How did that experience make you feel? How were your perspectives changed as a result of that experience?
3. How do you identify yourself in your career? Indicate it on scale of novice to master.
4. Suppose I were a new teacher in your school or district. What unspoken rules would I have to learn to follow in order to be accepted/stay out of trouble when it comes to teacher evaluation?
5. Describe an incident of teacher evaluation that has been powerful for you. How did it make you feel? How did you react when it happened? What lasting impact on you did it have?
6. What assumptions does the teacher evaluation policy make about teachers?
7. Suppose you were in a small group of fellow teachers at your grade level or in your department and someone shared that they got excellent marks on their teacher evaluation? How would you feel? How would you respond? How about if they got awful marks on their evaluation? How would you feel? How would you respond?
8. How does teacher evaluation impact your students and their families?
9. What is your school district currently doing to promote? What more should they be doing?
10. Please describe your own experience with teacher evaluation that has been meaningful to you. What have you learned from that connection? (If no example), what factors have prevented you from having a meaningful experience?
11. Who do you think benefits from teacher evaluation? How?

---

<sup>11</sup> Based on Lee Bell’s (2010) Interview Guide for the Storytelling Project as described in Bell, L. (2010). The storytelling project model: A theoretical framework for critical examination of racism through the arts. *Teachers College Record*, 112(9), 2295-2319.

12. What privileges have you experienced as a teacher in your school? What privileges have you been denied as teacher in your school?
13. How has teacher evaluation changed in your lifetime? Please give some examples of changes you have witnessed at a school, district, state and national level? Changes you have witnessed on a personal level?
14. How have your own attitudes/awareness about teacher evaluation changed over time? What brought about these changes?
15. What conditions do you think are necessary for teacher evaluation to improve?
16. How would teaching be different if teacher evaluation were eliminated?
17. What do you see as the most promising or hopeful actions being taken today in terms of teacher evaluation? If you had the power to change any of the things in the current teacher evaluation system, what would you change?
18. What else would you like to add to this discussion? What haven't I asked about that you think is important to mention?

Thank you again for your participation in the study, and as a reminder, all recordings and transcripts are accessible only to myself and pseudonyms will be assigned in any publications.

## **Appendix B: Teacher Inquiry Group Sessions**

**Protocol Title:** Teachers as Policy Actors: A Collaborative Teacher Research Project  
Exploring Teacher Evaluation in Flowertown

**Principal Investigator:** Filomena Hengst, Teachers College Doctoral Student  
[email address redacted], [phone number redacted]

### **Teacher Inquiry Group Sessions Protocol**

Based on Lee Bell's (2010) storytelling project as described in Bell, L. (2010). The storytelling project model: A theoretical framework for critical examination of racism through the arts. *Teachers College Record*, 112(9), 2295-2319. The following protocol outlines the facilitation of the Teacher Inquiry Group Sessions.

#### **Session 1 Facilitator Guide**

1. When participants enter, they will be asked to complete the Basic Storytelling Elements sheet. Prompt will read "Think about one of the first times you were evaluated as a teacher. Feel free to record your ideas if you would like to/refer to **Basic Storytelling Elements** checklist (see below), and we will go around the room and start to share our stories later in the session." Introduce each other by name.

##### **Basic elements of a story:**

- Setting (year, school, teaching assignment)
- Characters (persons involved)
- Plot (What happened?)

2. What makes you feel particularly pleased and successful as a teacher? Any ways these priorities may have changed over the years, for better or for worse?

3. What is teacher evaluation (policy)? Facilitate a discussion about what teacher evaluation is. Facilitate a discussion about what policy is. Offer a brief presentation about the history and structure of the teacher evaluation policy.

A. According to AchieveNJ, "'Evaluation" means an appraisal of an individual's professional performance in relation to his or her job description and professional standards and based on, when applicable, the individual's evaluation rubric."

B. What is the logic behind the teacher evaluation policy?

C. What is the purpose of this policy?

4. Display word clouds based on participants' 10 words. Next, we are going each develop and convey a story about one of your first experiences with teacher evaluation. Pretend you are telling it to a first-year teacher. Start with the phrase "I remember..." and don't worry about making the story perfect.



4. Debrief: Find commonalities among stories, reflect on individualistic experience and power structure. Display one of Cezanne's paintings of a tree and convey the idea that personal experience also manifests universal themes or ideas. Ask participants to draw a picture of the evaluation process and see if anything is realized from that process. Ask, what factors contribute to teacher evaluation policy? Create a poster to record the answers.
- 

### **Session 3 Facilitator guide**

1. Review poster from previous session.
2. Activity: Dramatize physical representations of concepts that underlie teacher evaluation. Each group will pick a concept related to teacher evaluation. They will develop their scenario, performs their embodied story/performance for the team using postures, gestures and movement to express concepts. Create a performative space where people can experience the familiar and unfamiliar in order to push their thinking. Groups of 2-3 persons will have 10 minutes to plan and 4 minutes to present each. If there are only 3 participants, we will individually come up with representations.
3. After pushing our thinking through embodied story/performance, we will again discuss the effects of the teacher evaluation system on our work, including an exploration of unintended and intended consequences and the factors that contribute to our characterization as such. Define "intended" consequences of the teacher evaluation system together and discuss. Do the same with "unintended" consequences. Discuss implications. Take notes on posters so all can track the discussion visually.
4. Offer examples of teachers working together to change policy.
5. What should we do? Co-author a design proposal/an action plan

### **Session 4 Facilitator guide**

#### **Presentation of findings/data and group analysis of findings**

1. Facilitator will place posters around the room with preliminary findings by the researcher.
2. Participants will go around with post-it notes (different colors per participant) to react, critique, and further analyze each of the findings. They can also write on the poster itself in their color marker.

Then, the group will discuss their data analysis comments together.

## **Appendix C: Verbal Recruitment Script**

Hi, I'm interested in having you participate in a teacher research project that I am conducting called "Teachers as Policy Actors: A Collaborative Teacher Research Project Exploring Teacher Evaluation in Flowertown." The expectations are that you would attend four teacher inquiry sessions and possibly a follow-up interview. During the teacher inquiry sessions, you will be involved in interactive dialog, storytelling activities and offering your opinions to a degree that you feel comfortable. We will work together to engage teacher evaluation policy and possibly come up with an action plan at the end. You would also agree to my observing you during a fourth month period and jotting anything that is said about teacher evaluation in conversation at meetings and gatherings. The time commitment totals 7-10 hours for the duration of the project from October to January. For the teacher inquiry group, snacks will be provided and you will receive credits toward a flex day for release from instructional duties in our school district. Please take a look at the consent forms, and let me know if you have any questions.

## **Appendix D: Informed Consent Form**

TEACHERS COLLEGE  
COLUMBIA UNIVERSITY  
INSTITUTIONAL REVIEW BOARD  
Box 151

### **TEACHERS COLLEGE INSTITUTIONAL REVIEW BOARD**

525 West 120th St. \* New York, NY 10027 \* 212-678-3000

[www.tc.columbia.edu/institutional-review-board](http://www.tc.columbia.edu/institutional-review-board)

**Protocol Title:** Teachers as Policy Actors: A Collaborative Teacher Research Project Exploring Teacher Evaluation in Flowertown

**Principal Investigator:** Filomena Hengst, Teachers College Doctoral Student  
[phone number redacted], [email address redacted]

### **INTRODUCTION**

You are being invited to participate in this research study called “Teachers as Policy Actors: A Collaborative Teacher Research Project Exploring Teacher Evaluation in Flowertown.” You may qualify to take part in this research study because **you are a K-5 elementary school teacher who identifies as female**. No more than 25 people will participate in this study and it will take 7-10 hours of your time to complete.

### **WHY IS THIS STUDY BEING DONE?**

The purpose of this qualitative dissertation study is to explore what teacher research project can reveal about teacher evaluation reform’s effects on teaching.

### **WHAT WILL I BE ASKED TO DO IF I AGREE TO TAKE PART IN THIS STUDY?**

If you decide to participate, you will be involved in interactive dialog, storytelling activities and offering your opinions about teacher evaluation to a degree that you feel

comfortable. This teacher inquiry group will be audio-recorded. If you do not wish to be audio-recorded, we can make arrangements for another researcher to take notes during the sessions with your consent. This will take place in the principal investigator's classroom in Flowertown, New Jersey.

You could then be asked to participate in a recorded interview with you where you will be asked to discuss the topic of teacher evaluation and your perspective on it further. This interview will be audio-recorded. After the audio-recording is written down the audio-recording will be deleted. If you do not wish to be audio-recorded, you will not be able to participate in the interview. This will take about 60-120 minutes. You will be given a pseudonym or false name/de-identified code in order to keep your identity confidential. This will take place at a location of your choice in a format that is face-to-face, by telephone or through online video chat at a time that is convenient to you.

You may also be observed by the lead researcher in your conversations during meetings and gatherings inside and outside work when discussing teacher evaluation for the duration of the study.

### **WHAT POSSIBLE RISKS OR DISCOMFORTS CAN I EXPECT FROM TAKING PART IN THIS STUDY?**

The harms or discomforts during the teacher inquiry group sessions or interviews that you may experience are not greater than you would ordinarily encounter in daily life while participating in discussions about yourself or teaching evaluation. However, there are risks to consider. Given that this is a study specifically for K-5 tenured, female teachings working in the same school district, material gathered from the group will be attributable to a relatively small number of individuals. Thus, it is possible that members of the school and district community may be able to infer participants' identities and/or surmise the source of some opinions contained in the presentations, reports or publications coming out of this work. **You do not have to answer any questions or divulge anything you don't want to talk, about and you can stop participating in the study at any time without penalty.** You will also be reminded of this risk should you say or recommend anything controversial. The principal investigator is taking precautions to keep your information confidential and prevent anyone from discovering or guessing your identity, such as using a pseudonym instead of your name and keeping all information on a password protected computer and in a locked file drawer. Additionally, the lead researcher is taking other measures such as recruiting a pool of people larger than those participating in the study, not revealing participants' names to anyone, relying on paraphrase rather than direct quotations, and asking participants to check information before it is included in any presentations, reports and publications.

### **WHAT POSSIBLE BENEFITS CAN I EXPECT FROM TAKING PART IN THIS STUDY?**



There is no direct benefit to you for participating in this study. Participants gain flex time and professional development hours in the district for participation in the study. Participation may benefit the field of teacher evaluation policy by offering teacher perspectives on the issue.

### **WILL I BE PAID FOR BEING IN THIS STUDY?**

You will not be paid to participate.

### **WHEN IS THE STUDY OVER? CAN I LEAVE THE STUDY BEFORE IT ENDS?**

The study is over when you have completed the interview, if selected. Otherwise, the study is over for you after participation in four sessions of the teacher inquiry group. However, you can leave the study at any time even if you haven't finished.

### **PROTECTION OF YOUR CONFIDENTIALITY**

The investigator will keep electronic or digital information (including audio recordings) stored on a computer that is password protected. What is on the audio-recording will be typed and the audio-recording will then be destroyed. There will be no record matching your real name with your pseudonym. Regulations require that research data be kept for at least three years.

Given that this is a study specifically of K-5 female elementary school teachers, interview material will be attributed to a relatively small group of individuals. Specific identification of any one person by colleagues, management or outside parties is unlikely, but possible. Preliminary results will be shared as a general report with the superintendent with no identifiers, and the final results of this study will be used for scholarly purposes only. Should an administrator inquire about information from the study, the lead researcher will not share anything without checking with the participants first.

For quality assurance, the study team, the study sponsor (grant agency), and/or members of the Teachers College Institutional Review Board (IRB) may review the data collected from you as part of this study. Otherwise, all information obtained from your participation in this study will be held strictly confidential and will be disclosed only with your permission or as required by U.S. or State law.

### **HOW WILL THE RESULTS BE USED?**

The results of this study will be published in journals and presented at academic conferences. Your name or any identifying information about you will not be published.

### **CONSENT FOR AUDIO RECORDING**

Audio recording is part of this research study. You can choose whether to give permission to be recorded. If you decide that you don't wish to be recorded, you will still be able to participate in this study.

\_\_\_\_\_ I give my consent to be recorded \_\_\_\_\_

Signature

\_\_\_\_\_ I **do not** consent to be recorded \_\_\_\_\_

Signature

### **WHO MAY VIEW MY PARTICIPATION IN THIS STUDY?**

\_\_\_\_ I consent to allow written, video and/or audio taped materials viewed at an educational setting or at a conference outside of Teachers College \_\_\_\_\_

Signature

\_\_\_\_ I **do not** consent to allow written, video and/or audio taped materials viewed outside of Teachers College Columbia University \_\_\_\_\_

Signature

### **WHO CAN ANSWER MY QUESTIONS ABOUT THIS STUDY?**

**If you have any questions about taking part in this research study, you should contact the principal investigator, Filomena Hengst at [phone number redacted] or [email redacted]. You can also contact faculty advisor Dr. Dirck Roosevelt at [email redacted].**

**If you have questions or concerns about your rights as a research subject, you should contact the Institutional Review Board (IRB) (the human research ethics committee) at 212-678-4105 or email IRB@tc.edu. Or you can write to the IRB at Teachers College, Columbia University, 525 W. 120<sup>th</sup> Street, New York, NY 1002. The IRB is the committee that oversees human research protection for Teachers College, Columbia University.**

---

### **PARTICIPANT'S RIGHTS**

- I have read and discussed the informed consent with the researcher. I have had ample opportunity to ask questions about the purposes, procedures, risks and benefits regarding this research study.
- I understand that my participation is voluntary. I may refuse to participate or withdraw participation at any time without penalty to future employment.
- The researcher may withdraw me from the research at his or her professional discretion, such as if I do not fit the selection criteria.

- If, during the course of the study, significant new information that has been developed becomes available which may relate to my willingness to continue my participation, the investigator will provide this information to me.
- Any information derived from the research study that personally identifies me will not be voluntarily released or disclosed without my separate consent, except as specifically required by law.
- Identifiers may be removed from the data. Deidentifiable data may be used for future research studies or distributed to another investigator for future research without additional informed consent from the subject or the representative.
- I should receive a copy of the Informed Consent document.

**My signature means that I agree to participate in this study**

**Print name:** \_\_\_\_\_ **Date:** \_\_\_\_\_

**Signature:** \_\_\_\_\_