Using Video to Reframe the Teacher Evaluation Process as a Professional Learning Experience: An Action Research Study

Andrew C. Matteo
USING VIDEO TO REFRAME THE TEACHER EVALUATION PROCESS AS A
PROFESSIONAL LEARNING EXPERIENCE: AN ACTION RESEARCH STUDY

A DISSERTATION

Submitted to the Faculty of
Montclair State University in partial fulfillment
of the requirements
for the degree of Doctor of Philosophy

by

ANDREW C. MATTEO
Montclair State University
Upper Montclair, NJ
May 2020

Dissertation Chair: Dr. Katrina Bulkley
MONTCLAIR STATE UNIVERSITY
THE GRADUATE SCHOOL
DISSERTATION APPROVAL

We hereby approve the Dissertation

USING VIDEO TO REFRAME THE TEACHER EVALUATION PROCESS AS A
PROFESSIONAL LEARNING EXPERIENCE: AN ACTION RESEARCH STUDY

of

Andrew C. Matteo
Candidate for the Degree:
Doctor of Philosophy

Graduate Program:
Teacher Education and Teacher Development
Certified by:

Dissertation Committee:

Dr. M. Scott Heimess
Vice Provost for Research and
Dean of The Graduate School

April 16, 2020

Date

Dr. Katrina Bulkley
Dissertation Chair

Dr. Kathryn Herr

Dr. Emily Klein
ABSTRACT

USING VIDEO TO REFRAME THE TEACHER EVALUATION PROCESS AS A PROFESSIONAL LEARNING EXPERIENCE: AN ACTION RESEARCH STUDY

by Andrew C. Matteo

The purpose of this action research study was to work with a group of teachers to refine an organizational routine dealing with teacher evaluation with the dual goals of (a) increasing the teachers’ sense of professionalism and improving the technical core of their work—instruction—and (b) moving toward a routine that the district will be able to use in the future. The research question that guided this work was, “How do we experience the organizational routine that I developed using video for teacher evaluation for improving teacher practice as compared to the existing teacher evaluation system, and how does our experience inform the routine’s continued development?”

This study takes place against the backdrop of over 100 years of American educational history in which teaching has not been treated as a profession as well as a more recent context in which teacher evaluation systems have focused on teacher accountability but failed to emphasize teacher development. As a result, our teacher evaluation system contributes little to the professionalization of teachers and does not align with elements of effective professional development to the improve teacher practice. Therefore, I set out to conduct a study that used the power of an organizational routine to potentially change teachers’ perceptions of being treated as professionals and to foster teacher learning and improve practice. The study took place over two cycles, with the teachers and me collaborating on revisions to the routine for the second of the two cycles.
Four findings emerged from this study. First, I explain how the participants viewed the existing evaluation system as transactional and deprofessionalizing. Second, I explore how both the first and second versions of the new routine increased the participants’ perceptions of being treated as professionals due to increased teacher agency in the teacher evaluation process. Third, I examine how the participants and I experienced improving teacher practice through the routine in light of the tenets of effective professional development. Finally, I examine the relational aspects of the study, specifically how relationships affected our experience and what lessons could be learned from the process.

This study has implications for teacher evaluation and administrators looking to conduct action research collaboratively with teachers around the use of organizational routines.

*Keywords:* teacher evaluation, video, organizational routine, professionalization, professional learning, professional development, teacher practice, action research
ACKNOWLEDGMENTS

This dissertation would not have been completed without the guidance of my advisor and dissertation chair, Dr. Katrina Bulkley. I am grateful for all that she has done to shepherd me to this point: joining me for countless Zoom calls (before the COVID-19 pandemic made them fashionable), putting up with my countless Google Docs, and providing constant and helpful feedback throughout this process—from coursework to proposal to dissertation to defense. Most importantly, I thank her for having confidence that I would cross the finish line.

I also thank the other members of my dissertation committee, Dr. Emily Klein and Dr. Kathryn Herr, for their excellent feedback on both my proposal and my dissertation. More importantly, I thank them for believing in me and pushing me to continually make this dissertation better.

Along with these professors, I acknowledge all the professors who have guided me through the Teacher Education and Teacher Development (TETD) program. I have learned important lessons from all of them, and I am a different person today than I was five years ago when I walked into University Hall. Thank you to Marcia Adirim, who was always so helpful navigating the logistics of the program and advocating on my behalf when it was needed.

I would be remiss if I did not offer sincere gratitude to the three teachers who served as my co-researchers in this process. They were always generous with their time and were fully committed to the work, and I cannot thank them enough for helping me explore this topic and achieve this goal.

I would not be here without my parents, Carl and Terri Matteo. My mother was my first teacher—and to this day is still my best teacher. She is an educator without equal, and I am grateful that she always takes an interest in my work and this research. I believe she will
actually read this dissertation from cover to cover. My father probably will not. But as Jim Valvano once said, “My father gave me the greatest gift anyone could give another person: he believed in me.” I can think of no better quote to express the gift that my father has given me through his belief in me. It is unwavering and unconditional. I am forever grateful for everything that my parents have done for me throughout my life. I will never be able to repay them. However, next time we go out to dinner, I’ll make the reservations.

Finally, I need to thank Heather, my critical friend throughout this dissertation. It is a role she was born to play. It also happens that she is my wife. And fellow TETD student. And now fellow doctor! I am in awe of her dedication, attention to detail, and work ethic. She completed her own dissertation while being a great mother to our young daughter, Grace, and while pregnant with our now infant daughter, Ella. Along with completing her own dissertation, she helped me with mine in countless ways. I would not have accomplished this without her love and support—and critiques.
DEDICATION

This dissertation is dedicated to my two amazing daughters, Grace and Ella. I hope your “dada” has made you proud. Every day I am proud of you and cannot wait to see all that you accomplish in this life. I have no doubt that you will make your mark and make this world a better, kinder, and more just place.
CONTENTS

CHAPTER 1: INTRODUCTION ................................................................................................... 1

Problem Statement .............................................................................................................. 2

Background for the Study ................................................................................................... 7

Purpose of the Study and Research Question ..................................................................... 9

CHAPTER 2: REVIEW OF THE LITERATURE ....................................................................... 10

Teaching as a Profession ................................................................................................... 10

Teacher Evaluation ........................................................................................................... 13

Professional Development ............................................................................................... 19

Organizational Routines .................................................................................................. 30

CHAPTER 3: RESEARCH DESIGN AND METHODS ............................................................. 38

Methodological Approach: Qualitative Action Research ............................................... 38

Context .............................................................................................................................. 39

Ethics and Positionality .................................................................................................... 40

Participants ......................................................................................................................... 41

The Routines ....................................................................................................................... 44

Data Collection .................................................................................................................. 46

Data Analysis ...................................................................................................................... 49

CHAPTER 4: FINDINGS CHAPTER ......................................................................................... 51

The Routines ....................................................................................................................... 52
LIST OF TABLES

Table 4.1 Comparison of First and Second Versions of Routine ........................................ 57
LIST OF FIGURES

Figure 2.1 Conceptual framework for my action research study................................. 36
Figure 4.1 Example of linking feedback to multiple teacher performance standards. .............. 94
Figure 4.2 Example of how a teacher can find time-stamped comment on Swivl Cloud.......... 96
Figure 4.3 Example of time-stamped comment from administrator in Swivl Cloud............. 97
Figure 4.4 Example of comment with linked resource................................................. 112
CHAPTER 1: INTRODUCTION

For decades, professional athletes have been reviewing game footage on video to help them reflect on and improve their performance. With their coaches, they identify strengths and highlight opportunities for improvement. However, teachers almost never get a chance to see themselves carrying out their craft or receive feedback from their supervisors that is tied to specific actions or teaching decisions they made and can view. It is hard to improve your practice if you do not have the opportunity to view yourself performing and if the feedback you receive is not specific, targeted, and tied to specific actions and decisions.

Some teachers have used video to advance their careers, but use of video in teacher evaluation remains uncommon. For instance, some preservice teachers have used video to support applications for teaching positions under the Educative Teacher Performance Assessment (edTPA), and a few have used video to enhance their applications for certification by the National Board for Professional Teaching Standards. However, most teachers have no opportunity to see themselves teach or to receive from an administrator feedback directly tied to video evidence of their teaching. Being aware of the complexity of teaching, the messiness of classroom instruction, and the need to focus on instructional decisions while teaching, I wondered whether using video would help teachers break down the complexity of teaching into more manageable chunks (Klein & Taylor, 2017; Sherin, Linsenmeier, & van Es, 2009).

The vast majority of evaluation is done via traditional in-class observations in which the supervisor visits the teacher’s classroom and writes up a report about what they observed, and then the two meet for a postobservation conference to discuss what they recall about the teaching. I designed an organizational routine that is more inviting and, I hope, more effective than traditional observations in helping teachers improve their practice. For the purposes of this
dissertation, I used Feldman and Pentland’s (2003) definition of organizational routine: “a repetitive, recognizable pattern of interdependent actions, involving multiple actors” (p. 96). The organizational routine consisted of having teachers videotape themselves, choose a lesson that they would like feedback on, complete a self-reflection, share their lesson with me via an online platform, and then receive feedback directly tied to the video followed by a debrief. I designed this organizational routine to be carried out in the district where I currently serve as the director of curriculum. I envisioned conducting two cycles of this action research with several teachers from across the district, with the goal of helping teachers reflect more about their practice, giving them more ownership over the evaluation process, and allowing the observer to provide better feedback tied to specific actions that could be seen and heard in the video.

**Problem Statement**

Due to a variety of factors, teaching has never managed to break through as a true profession on par with medicine or law (Darling-Hammond & Hyler, 2013; Mehta, 2013). Although teaching should be considered a profession because it is nonroutine work that requires skill and expertise, the way that it is structured does not meet Mehta’s (2013) definition of a profession, which is “characterized by deep levels of knowledge and expertise, professionally shared standards of appropriate practice, and the use of judgment and discretion in applying that common knowledge to particular situations” (p. 469). Rather, due to historical and sociological factors, teaching is set up as a loosely coupled bureaucracy. As Mehta explains,

> The more general historical pattern in American teaching . . . has been a loosely coupled bureaucracy that is hierarchical in its distribution of authority but fairly weak in its specification of the actual procedures that govern practice. District and, later, state and federal officials prescribe what has to be learned, at what pace, and with what
consequences, but then teachers are given considerable freedom to decide how to reach these ends. In this arrangement, creating a common technical core that might reliably produce such outcomes is not really anyone’s responsibility. (p. 468–469)

It has long been argued that education is full of loosely coupled systems (Weick, 1976). In defining this idea, Weick writes, “By loose coupling, the author intends to convey the image that coupled events are responsive, but that each event also preserves its own identity and some evidence of its physical or logical separateness” (p. 4). In other words, to say that items in a system are loosely coupled is to say that the connection between one level of the system (i.e., school or district administration) and another (classroom practice) is not tight and causal. Therefore, it is very difficult in loosely coupled systems for one component of the system to control or substantially influence another. To say that schools are systems of this sort is to say that it is very difficult for external forces or pressure to significantly affect classrooms due to the relatively high autonomy that characterizes classrooms. Loose coupling has advantages and disadvantages. An advantage is that it allows for local adaptations, which are important for teacher agency, autonomy, and innovation. On the negative side, loose coupling makes it harder to achieve any systemic change at scale and is often cited as a reason that educational improvement has been so elusive (Coburn, 2004; Spillane, Parise, & Sherer, 2011; Weick, 1976).

The bureaucratic structures and loose coupling of schools have had profound implications for many of the routines involved in the work of schooling, including evaluation. Evaluation, one of the most significant formalized routines carried out in schools involving instruction, has two main purposes—accountability and teacher development (Darling-Hammond, Wise, & Pease, 1983; Woulfin, Donaldson, & Gonzales, 2016). Although the pendulum has swung back and forth between these two purposes over the years, throughout the history of American
education and especially recently and up to the present, accountability—and trying to weed out the “bad” teachers—appears to have been the dominating factor in the design of most teacher evaluation systems (Marzano, Frontier, & Livingston, 2011; Mehta, 2013; Papay, 2012). Papay (2012) argues that this is a problem, for if we aim to help students learn better through our teaching, then we have to refocus our evaluation on teacher development. We will be better served, he maintains, if we “shift the debate from an argument about the best way to identify the top and bottom performers to a discussion of how best to use performance evaluation to improve teacher development and boost student learning in schools” (p. 125).

Since 2009, 46 states have reformed their teacher evaluation process (Donaldson, 2016). Consequently, all 50 states now have policies for standards-based teacher evaluation (Stronge and Associates, 2018). Advocates argue that because it provides a common language about instruction and focuses on evidence-based feedback, standards-based teacher evaluation is more conducive to facilitating teacher growth than traditional drive-by evaluations of the past (Donaldson, 2016; Papay, 2012; Toch and Rothman, 2008). In New Jersey, all districts are now required, under a policy called AchieveNJ, to implement standards-based evaluations aligned to a standards-based teacher practice approved by the New Jersey Department of Education (NJDOE). This requirement was enacted through a 2012 law known as the Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) Act, which resulted in AchieveNJ. For example, in our district we follow the Stronge model, which is based on James Stronge’s framework of effective teaching (Stronge, 2018).

The Stronge model, officially called the Stronge Teacher Effectiveness Performance Evaluation System (TEPES), is organized into seven performance standards that define a teacher’s practice and are rooted in research on teaching. Those seven performance standards
are professional knowledge, instructional planning, instructional delivery, assessment of for student learning, learning environment, professionalism, and student progress. For each of those seven performance standards Stronge lists multiple performance indicators that more specifically outline what effective practice looks like in that standard (Stronge and Assoc., 2016b). Our district selected this model because it is grounded in research and flexible. The system promotes itself as “a balance between structure and flexibility. It is prescriptive in that it defines common purposes and expectations, thereby guiding effective practice. At the same time, it provides flexibility, thereby allowing for creativity and individual teacher initiative” (Stronge and Assoc., 2016b). An example of this flexibility is that the performance indicators can be modified by individual school districts to align with district goals and the local context. We take advantage of this flexibility by reviewing and revising the performance indicators each spring with a diverse group of stakeholders including teachers.

The Stronge model uses multiple data sources, including observations, a documentation log, and a measure of student progress. The documentation log allows the teacher to submit artifacts and examples of work that show evidence of effectiveness in a particular standard. In this way the teacher is supposed to be able to contribute to the overall portfolio that will be used as the basis for their yearly evaluation. Another reason we chose the Stronge model is that the observer does not provide a rating after an observation. This is different from some other evaluation systems, including the Danielson framework, which require rating at each and every observation. This was important because we wanted the observation to be seen as a formative experience focused on providing rich feedback to promote growth. At the end of each school year the teacher is evaluated in each performance standard using a 4-point rubric with the
following categories: highly effective (4), effective (3), partially effective (2), and ineffective (1). The teacher then receives an overall score using the same categories.

The move toward standards-based teacher evaluation has altered the organizational routine of teacher evaluation. Within the overall idea of organizational routine, it is important to differentiate between the ostensive routine and the performative routine (Feldman & Pentland, 2003). The ostensive routine is the routine in its idealized form, whereas the performative routine is the routine as practiced. When thought of within the context of teacher evaluation, the ostensive routine is how teacher evaluation is supposed to be carried out according to the design put forth in AchieveNJ, whereas the performative is the act of teacher evaluation as implemented locally.

The focus on accountability has led to teacher evaluation systems that, while paying lip service to teacher development in their missions, often result in practices that do very little to help teachers develop (Donaldson and Donaldson, 2012). Furthermore, many factors affect how these standards-based evaluations are implemented in a district and school (Donaldson and Donaldson, 2012; Stronge, 2018; Woulfin et al., 2016). Two big factors at the district level are how district leaders interpret and frame the policy for their district and train principals to carry out the teacher evaluation process and support them as they do so (Steinberg and Sartain, 2015; Woulfin et al., 2016). At the school level, Steinberg and Sartain (2015) found that the effectiveness of teacher evaluation depends on the principal’s capacity to provide targeted instructional guidance and the teacher’s ability to receive and respond to instructional feedback in a manner that improves student achievement.

For many school districts, including my own, supervisors and building principals do not have enough time to observe teachers and give them meaningful feedback (Woulfin et al., 2016).
This is a problem because “providing effective feedback is the bridge that connects teacher evaluation to improvement” (Stronge & Associates, 2016b, p. 57). Even when teachers do receive feedback, they don’t get to “see” what the supervisor is talking about. They usually have to rely on remembering. Without a shared understanding of the events that transpired during the lesson, teachers often question feedback from observers because they do not agree with the recollection of the lesson (Center for Education Policy Research [CEPR], 2014). The postconference requires both participants to remember events from previous days, and sometimes those recollections differ (CEPR, 2014).

Stronge and Associates (2016b) recognizes the importance of being descriptive when it writes, “Evaluators are not expected to describe everything, but the more descriptive they are, the more beneficial the feedback will be” (p. 57). However, it is very difficult for an observer to provide specific, descriptive, and evidence-based feedback that is anchored in what actually happened. I contend that the use of video may minimize the need for the observer to describe the teaching, because the video recording will capture many of the data of the teaching. Therefore, the observer can focus on connecting what he or she sees to a performance standard and providing feedback. Thus, the evaluation process will contribute to the professionalization of teaching by helping teachers improve their skills in an agreed-upon body of knowledge. As a member of the profession, the teacher will play an active role in the work.

**Background for the Study**

I began as my district’s director of curriculum, instruction, and assessment in July 2017. In this role, I coordinated the district’s teacher evaluation and professional development programs. In the spring of 2018, the superintendent in my district, my direct supervisor, asked me to oversee a pilot program of using video as a tool for coaching. He had recently attended a
professional development workshop at the Harvard Graduate School of Education, and this sparked his interest in bringing the practice of using video for professional learning into our district. Specifically, he wanted me to organize and oversee a trial of what it would look like if a group of our teachers shared videos and received feedback from our district’s literacy consultant. Drawing on the research and literature from the conference at Harvard, he shared several reasons why this was a worthwhile endeavor.

As part of my action research course at Montclair State University in spring 2018, I conducted an action research project for the semester with the possibility that it could evolve into a dissertation topic. The action research project consisted of using video for self-reflection, peer collaboration, and video coaching. Using the literature from the Harvard conference as my guide, I also explored other possibilities, including using video to build a library of best practices and using video in the formal evaluation cycle. As I read more about teacher evaluation and the possible role of video in that process, I became interested in conducting action research on teachers’ perceptions of the teacher evaluation process by engaging in a new organizational routine—the use of video for classroom observations.

Another reason I decided to focus this dissertation on teacher evaluation is the central role that it has played in the educational landscape over the last six years. The shift to required, standards-based observations has been one of the biggest changes, if not the biggest change, to day-to-day teaching and learning in our schools. There has also been a recent push to have building administrators serve more as instructional leaders. With the implementation of AchieveNJ, principals now must spend a significant part of their day observing teachers. However, this is often difficult due to the leader’s hectic schedule and competing demands on their time. Although getting principals into classrooms observing instruction and then discussing
it with teachers is good, it is important that this work be designed and organized well so that it ultimately helps teachers improve their practice. If designed and implemented poorly, observation may not only waste the time of those involved; it may also fail to improve instruction and may further depprofessionalize teaching.

**Purpose of the Study and Research Question**

As director of curriculum in my district, I designed a new organizational routine around teacher evaluation and invited three veteran teachers to help me refine it. The purpose of this action research study was to work with those teachers to refine the routine, with the multiple goals of increasing the teachers’ sense of professionalism, improving the technical core of their work — instruction—and developing a routine that the district could use in the future. For this action research study, I asked, “How do we experience the organizational routine that I developed using video for teacher evaluation for improving teacher practice as compared to the existing teacher evaluation system, and how does our experience inform the routine’s continued development?”

The teachers were asked to describe their experience with using video for teacher observation and to specify how, if at all, it shaped their professional practice. I hoped that this work would translate into a more effective approach to teacher evaluation that focuses more on professional development than accountability. Also, as I, being a participant in this action research study, coconstructed the second cycle with my participants, I examined what I had already learned about using video for teacher observation and how the process of the study shaped the routine.
CHAPTER 2: REVIEW OF THE LITERATURE

This chapter begins with an overview of teaching as a profession and how that connects to the history and evolution of teacher evaluation. The purpose of these historical overviews is to demonstrate how teacher professionalism and teacher evaluation are linked and how they both provide the context for the current issues with teacher evaluation. Due to the history of teaching in the United States, the routine of teacher evaluation has been carried mostly bureaucratically, focused on accountability to the detriment of teacher development (Marzano et al., 2011; Mehta, 2013). From there, I turn to the different conceptualizations of teacher learning before describing the tenets of effective professional development as outlined by Darling-Hammond, Hyler, and Gardner (2017). Next, I review the existing literature on using video as a professional development tool and how that intersects with ideas of effective professional development. I make the case that teacher evaluation can and should be reframed with an organizational routine that includes some of the characteristics of effective professional development and that has potential to bring about endogenous change, or “change that comes from within organizational routine: change that is a result of engagement in the routine itself” (Feldman & Pentland, 2003, p. 19). As part of this argument, I review the literature on organizational routines before ending this chapter with the conceptual framework underpinning my study, which includes the major concepts mentioned above.

Teaching as a Profession

The legacy of teaching in the United States contributes to many of the current issues facing teaching today as it continues to attempt to be defined as a true profession with a systematic method for translating the knowledge base of the profession into practice (Mehta, 2013). Beginning with the one-room schoolhouses of the Colonial period, teachers required no
specialized teacher training, operated in isolation, and rarely interacted with other adults (Goldstein, 2014; Lortie, 1975). When increased urbanization of the population in the early 19th century led to the creation of larger schools in the U.S., teachers in these emerging towns increasingly came together under one roof in a single school, but remained isolated (Lortie, 1975). For the most part, each teacher operated as an independent contractor or “separate cell” within the school, with no shared practices or any specialized training (Lortie, 1975). In the late 19th century, John Dewey proposed the idea of laboratory schools in which teachers served as researchers of their own practice (Dewey, 1896; Mehta, 2013). This was an early idea of what teaching might look like if it were treated as a true “profession” that involved its practitioners in developing and implementing knowledge (Mehta, 2013).

Using Mehta’s (2013) argument, teaching is more appropriately understood as a profession due to the high levels of nonroutine work requiring high levels of knowledge, skill, and expertise. It is important to think of teaching as a profession because with that designation come shared standards of practice along with common knowledge that can be shared across contexts. These shared standards and common knowledge would allow teaching to be deprivatized and for best practices to emerge. Mehta argued:

Evident in Dewey’s ideas is an early form of what it might look like to build a profession based on practical knowledge: teachers and researchers working alongside one another to produce new knowledge; knowledge that is organized in part around practical questions coming from the field rather than entirely out of the heads of university researchers; a healthy, interdependent relationship between basic and applied knowledge; and teachers treated as professionals capable of taking “intellectual initiative” and organizing their own work rather than implementing the ideas of others. (p. 471)
However, Dewey’s view of teaching grounded in practice-based research largely lost out to the view proposed by Edward Thorndike and supported by higher education faculty that research and knowledge should be produced by researchers at universities and then ultimately implemented by teachers in the classroom (Mehta, 2013). This separation of knowledge from practice would have profound implications for the organization of teaching as a bureaucracy rather than a profession (Mehta, 2013). Two of these implications are the focus on accountability in evaluation and a lack of practice-based professional development for teachers.

For the purposes of this study, I define practice-based professional development as any formal professional learning activity that takes place in the actual context of teaching (Ball & Cohen, 1999; Cochran-Smith & Lytle, 1999; Lampert, 2010). When knowledge and practice are separated as they have traditionally been in education, there is no need to even attempt to organize and facilitate practice-based professional development. If knowledge must be formed outside the classroom, then it makes no sense to look inside the classroom for professional development opportunities. In other words, if knowledge must be developed by researchers and academics, then the classroom is merely the vehicle for implementing that knowledge. For example, traditional professional development has occurred in “one-off” workshops conducted by outside experts or through teachers taking graduate courses at local institutions of higher education. However, in the last few decades, professional learning has been becoming more practice based (Ball & Cohen, 1998; Cochran-Smith & Lytle, 1999; Lampert, 2009). This includes opportunities for teachers to study their own practice and that of their colleagues, including through the use of video.
Teacher Evaluation

The outcome of the Dewey-Thorndike battle had implications for the view of teaching as a profession and, in turn, for how teaching was and (it was thought) should be evaluated. Instead of being considered a full profession, teaching took on many of the characteristics of a bureaucracy, which led to an evaluation system focused on managerial control and accountability (Mehta, 2013).

History of Teacher Evaluation

As described in the problem statement, teacher evaluation has two, often competing, purposes: accountability and teacher development. Here I overview the history of teacher evaluation in the United States and explore how those two purposes have impacted how teacher evaluation is carried out as a routine in American schools.

1700s to mid-1800s. Throughout the 18th century and well into the 19th, education was not seen as a field of expertise (Marzano et al., 2011). The teacher was often hired and supervised by the local government or clergy, who had no expertise in pedagogy or subject matter knowledge. Therefore, supervisory practices and feedback provided to teachers varied greatly. In the mid-1800s, teaching came to be seen as a more complex task, which therefore required better feedback in order for the teacher to develop expertise. As schools became more complex, there arose “principal” teachers whose job was to provide specialized guidance for teachers. Although teaching did begin to be viewed as a more complex task that required certain skills, there was little discussion about what these skills actually were.

Late 1800s to World War II. Around the turn of the 20th century, two major views of education competed—the progressive view championed by John Dewey and the scientific view championed by Edward Thorndike (Mehta, 2013). As described above, Thorndike’s approach
largely won out, and along with it came the scientific view of management proposed by
Frederick Taylor (Marzano et al., 2011). This approach viewed schools as factories with
teachers as factory workers whose work should be evaluated and measured accordingly. This led
to a focus on measurement to make sure that teachers and schools were effective and efficient.

Post WWII. After World War II, society began to view supervision of teachers more
developmentally, and classroom observation became an important part of a supervisor’s job.
This set the stage for clinical supervision, which was invented by Morris Cogan at Harvard
University in the mid-1950s (Marzano et al., 2011; Pajak, 2008). Although the idea of
preobservation, observation, and conference was known to exist before explicit ideas of clinical
supervision, Cogan’s model was more systematic and refined than anything that came before it
(Marzano et al., 2011; Pajak, 2008). Furthermore, it focused on the teacher as an active
participant in their own supervision (Zepeda, 2007). Cogan’s model was aligned with the view
of the teacher and teaching proposed by Dewey and the developmental purpose of education.
Cogan (1973) put it this way:

A cornerstone of the supervisor’s work with the teacher is the assumption that clinical
supervision constitutes a continuation of the teacher’s professional education. This does
not mean that the teacher is “in training,” as is sometimes said of preservice programs. It
means that he is continuously engaged in improving his practice, as is required of all
professionals. In this sense, the teacher involved in clinical supervision must be
perceived as a practitioner fulfilling one of the first requirements of a professional—
maintaining and developing his competence. He must not be treated as a person being
rescued from ineptitude, saved from incompetence, or supported in his stumblings. He
must perceive himself to be engaged in the supervisory processes as a professional who continues his education and enlarges his competences. (p. 21)

One of Cogan’s students, Robert Goldhammer, continued to evolve the model, ultimately publishing a book in 1969 that laid out a five-phase model of clinical supervision: preobservation conference, classroom observation, analysis, supervision conference, analysis of the analysis (Marzano et al., 2011). Although clinical supervision became widely used in education, its implementation (performative) often betrayed its original intent (ostensive). Originally designed as a way to encourage rich, trusting dialogue between teacher and supervisor for the purposes of reflection and growth, the process ultimately was reduced to a step-by-step procedure that should be strictly followed (Marzano et al., 2011).

1960s–1970s. The next major shift in instructional supervision and teacher evaluation came through the work of Madeline Hunter. Hunter is best known for her structured seven-step framework for a teaching lesson, which soon became commonly used in lesson design in the United States and became the substance of the preconference, observation, and postconference that made up the clinical supervision model (Marzano et al., 2011). However, Hunter contributed other important ideas to the field of teacher evaluation besides her lesson format. These contributions included proposing a common language of instruction to help improve professional development and script taping. In script taping, the supervisor would write down what he or she saw in the classroom and then categorize those actions by whether or not they promoted learning.

1980s. After the domination of the rigid approaches to supervision that were promoted through clinical supervision and the Hunter lesson structure, the early 1980s saw the rise of another view of supervision that called for more developmental, differentiated approaches.
These approaches emphasized the overarching goal of improving instruction and focused on providing different models of supervision based on the individual teacher being supervised. Within this context of competing approaches to supervision, the RAND group conducted a study to determine what types of teacher evaluation practices were actually occurring across the United States. The results of the study were published in a report entitled *Teacher Evaluation: A Study of Effective Practices* (Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984). The report found that, overall, evaluation practices were very “formulaic” and more consistent with the more rigid approach to supervision. An interesting conclusion of the report was that although some developmental approaches were in place, those systems were not specific enough to improve instruction. The report found that teachers themselves were the strongest supporters for a more standardized approach to supervision. Wise et al. (1984) reported that, “in their view, narrative evaluation provided insufficient information about the standards and criteria against which teachers were evaluated and resulted in inconsistent ratings among schools” (p. 16). The RAND report also identified four common issues with supervision and teacher evaluation, including lack of principal ability to conduct evaluations, teacher resistance to feedback, a lack of consistent evaluation practices based on teacher competencies, and lack of training for those conducting the evaluations (Wise et al., 1984).

**1990s.** Although Madeline Hunter proposed the idea of a common language for instruction, the RAND study found that this idea was not being enacted consistently or to scale. It was in this reality that, in 1996, Charlotte Danielson published her book *Enhancing Professional Practice: A Framework for Teaching*, which became the basis for the Danielson model, which was seen as a more comprehensive framework of teaching (Danielson, 1996). The Danielson framework, or Danielson model, identified 76 elements of quality teaching housed
within four domains: planning and preparation, the classroom environment, instruction, and professional responsibilities (Danielson, 2007). Each of these 76 elements was then broken down into four levels of performance (unsatisfactory, basic, proficient, and distinguished). This could be seen as the beginning of the standards-based teacher evaluation movement.

2000s–2010s. Although Danielson can be seen as the beginning of the standards-based teacher evaluation movement, in the first decade of the 21st century student achievement was increasingly used in teacher evaluation. The passage of the federal No Child Left Behind Act of 2001 (NCLB) and the establishment in 2009 of the federal Race to the Top (RttT) fund swung the pendulum toward an increasing focus on accountability in teacher evaluation (Donaldson and Donaldson, 2012; Woulfin et al., 2016). Many of these teacher evaluation systems attempted to achieve both goals of teacher evaluation by including student achievement data to hold teachers accountable for their performance and including a standards-based evaluation system that, while also holding teachers accountable to standards, was more aligned with the teacher development goal (Donaldson, 2016). In the next section, I provide a brief overview of one of those systems—AchieveNJ—that was implemented and still remains in effect in New Jersey.

AchieveNJ

As noted, the end of the first decade of the 21st century saw a focus on teacher evaluation across the country, including New Jersey. In 2010, the governor of New Jersey, Chris Christie, convened the New Jersey Educator Effectiveness Task Force, which was charged with examining the current teacher evaluation system. The task force released a report in March 2011 that listed several ways to improve the current system, including basing the evaluation system on clearly defined and high standards, making sure the new program is implemented with fidelity to ensure validity and reliability, and being consistent across school districts (NJDOE, 2014a). Out
of this work came a statewide pilot program to test out some of these proposals. As a result of this pilot program, in 2012 the state legislature passed and Governor Christie signed into law the Teacher Effectiveness and Accountability for the Children of New Jersey (TEACHNJ) Act. TEACHNJ laid out the new requirements for teacher evaluation in the state (NJDOE, 2014a). This law was then translated into a new teacher evaluation system known as AchieveNJ that was first implemented in the 2013–2014 school year.

The new law required teacher effectiveness to be determined at least partly by student achievement. According to TEACHNJ, “Changing the current evaluation system to focus on improved student outcomes, including objective measures of student growth, is critical to improving teacher effectiveness” (p. 1). The new law also required all school districts to implement a teacher evaluation system based on an approved teacher-practice instrument aligned to the New Jersey Professional Standards for Teachers.

**Standards-Based Teacher Evaluation**

Observations have been found very useful in improving instruction when the information provided to the teacher through the observation is both accurate and useful (Hill & Grossman, 2013). One aspect of this most recent teacher evaluation reform movement has been the common language of instruction promoted through the use of a standards-based framework along with a focus on evidence-based feedback within that framework (Stronge & Associates, 2018), which can be seen as part of a movement to reemphasize the teacher development part of evaluation and to reframe teacher evaluation as a formative assessment process. One proponent of this reframing has been John Papay (2012), who argued that “if teacher evaluation is to improve student learning systematically, it must be used as a tool to promote continued teacher development” (p. 124). He proposed a reframing of teacher evaluation from measurement to a
professional development tool with the goal of improving “instruction by developing teachers’ instructional capacity and effectiveness. . . . The evaluation system can and should be seen as a professional development tool and should be evaluated on its ability to raise instructional proficiency and student learning” (133). Toch and Rothman (2008) agreed that standards-based teacher evaluation is more conducive to teacher improvement than the drive-by evaluations that were common before this. In my district, we have adopted and use the Stronge Teacher Effectiveness Performance Evaluation System, one of the teacher evaluation rubrics approved by the state of New Jersey. In the next section, I overview this system and its underlying theory of effective teaching.

**Stronge Teacher Effectiveness Performance Evaluation System**

As described earlier, the Stronge Teacher Effectiveness Performance Evaluation System (TEPES) uses multiple data points to determine teacher effectiveness, including formal observations. For formal observations, evaluators observe a lesson and then provide feedback based on performance standards and indicators. The Stronge system is based on the premise that “providing effective feedback is the bridge that connects teacher evaluation to improvement” (Stronge & Associates, 2b p. 57). For feedback to be effective it must be evidence based. In one of their fact sheets, Stronge & Associates (2016b) explains, “The evidence collected should be as specific and objective as possible and should focus on the performance standards and indicators. . . . Evaluators are not expected to describe everything, but the more descriptive they are, the more beneficial the feedback will be” (p. 57). This view of supervision and evaluation lends itself nicely to the use of video, because both the teacher and the evaluator can view the lesson with their own eyes and thereby more accurately describe it.

**Professional Development**
Multiple scholars (DeSimone, Porter, Garet, Yoon, & Birman, 2002; Hawley & Valli, 1999) found that teacher learning and student achievement are tightly linked, suggesting that the improvement of schools rests on improving teaching. Furthermore, Hawley and Valli (1999) found that as teaching by telling is replaced by teaching for understanding, with a focus on mastery of complex knowledge and problem solving, this more complex type of teaching requires more advanced teaching skills, which in turn require better professional development for teachers. In other words, teacher (or professional) learning should be an outcome of effective professional development. I begin this section about professional development with an overview of the three major conceptions of teacher learning.

**Conceptions of Teacher Learning**

This brief discussion is organized into three subsections that represent the three major popular conceptions of teacher learning: knowledge-for-practice, knowledge-in-practice, and knowledge-from-practice (Cochran-Smith & Lytle, 1999; Lampert, 2010). These three different conceptions of teacher learning have led to very different structures of professional learning in the field.

**Knowledge-for-practice.** Due to the separation of knowledge and practice as outlined by the organizational structures that were in place, the dominant view of teacher learning during the 20th century was one that involved knowledge-for-practice (Cochran-Smith & Lytle, 1999). In this view of teacher learning, the teacher was the implementer of another person’s knowledge. They themselves were not responsible for constructing knowledge, but were simply the deliverers of information. From this perspective, teachers were encouraged to be well versed in the knowledge base of their content area so that they could transmit that knowledge base to their students (Cochran-Smith & Lytle, 1999). This belief led to a particular type of professional
development that is often conducted by experts in one-off workshops that regularly do not involve the teachers but are, rather, something that is done to the teachers (Ball & Cohen, 1999). Scholars argue that this type of learning is insufficient for teachers, because it does not take place in the “unpredictable situations we call classrooms” and therefore teachers are often unable to transfer their theoretical learning into practice (Ball & Cohen, 1999, p. 10). In fact, Shulman (2004) argues that teaching is in fact more complex than the established profession of practicing medicine: “The only time a physician could possibly encounter a situation of comparable complexity would be in the emergency room of a hospital during or after a natural disaster” (p. 258).

**Knowledge-in-practice.** This next conception of teacher learning focuses on the more practical aspects of teaching and considers the decisions that a teacher makes in the classroom setting (Cochran-Smith & Lytle, 1999). In this view, teachers need to be able to make quick judgments and respond to situations in the moment. Ball and Cohen (1999) argue that the type of knowledge needed for this kind of decision-making is best learned through experience and, in fact, must be learned through the practice of teaching. While one-day workshops suffice for the knowledge-for-practice view of teacher learning, this knowledge-in-practice view calls for “facilitated teacher groups, dyads composed of more and less experienced teachers, teacher communities, and other kinds of collaborative arrangements” (Cochran-Smith & Lytle, 1999, p. 263). Moreover, this conception breaks from the long tradition that knowledge comes from outside experts by assuming that “professional expertise comes in great part from inside the teacher profession itself” (Cochran-Smith & Lytle, 1999, p. 267).

**Knowledge-of-practice.** Similar to the second conceptualization, knowledge-of-practice breaks sharply with the knowledge-for-practice view of teacher learning, which draws a clear
distinction between knowledge and practice (Cochran-Smith & Lytle, 1999). This conceptualization of teacher learning is grounded in the belief that teachers must be constantly learning, studying their own practice, and developing a stance of inquiry (Ball & Cohen, 1999; Cochran-Smith & Lytle, 1999). This view of teacher learning studies the practice of teaching, “where inquiry is regarded as part of large efforts to transform teaching, learning, and schooling (Cochran-Smith and Lytle, 1999, p. 278).

I believe that the teacher evaluation process provides a great opportunity for teachers and administrators to coparticipate in rich practice-based professional development that comes out of the knowledge-in-practice and knowledge-of-practice conceptualizations. In order for them to do this, the teacher evaluation process must be rooted in these conceptualizations and must be planned and enacted according to the principles of effective professional development, which are described below.

**Effective Professional Development**

Although it may be clear that professional development matters, it has been less clear what constitutes effective professional development. For this section, I will use a comprehensive review of effective professional development (PD) in which Darling-Hammond et al. (2017) reviewed 35 studies that “have demonstrated a positive link between professional development, teaching practices, and student outcomes” (p. v) while supporting their argument with several seminal pieces on professional development.

According to Darling-Hammond et al.’s 2017 review, effective professional development is “structured professional learning that results in changes in teacher practices and improvements in student learning outcomes” (p. 2), because it (a) provides feedback and reflection, (b) involves
active learning, (c) is content focused, (d) includes coaching and support, (e) is collaborative, (f) uses modeling, and (g) is sustained.

In the following sections I provide more information about each of these seven elements, using several seminal pieces on professional development to provide a broader research base for the various characteristics, and then make explicit connections to how the first four characteristics can be found in my routine, when appropriate. As a reminder, the routine that I have designed consists of having teachers videotape themselves, choose a lesson that they would like feedback on, complete a self-reflection, share their lesson with their observer (me), and then receive feedback directly tied to the video, followed by a debrief.

**Feedback and reflection.** Darling-Hammond and her colleagues (2017) found that effective professional development usually provides time for teachers to receive feedback, reflect on their practice, and then effect the changes that their reflection led them to commit to. Hawley and Valli (1999) also found that the feedback in effective professional development must be information rich by incorporating multiple sources of information on outcomes for students based on experience and data. The feedback and reflection component is strongly aligned with the use of video for teacher evaluation. By using video and time-stamped feedback, evaluators can effectively pursue a major goal of teacher evaluation: improving the specificity and quality of feedback. Furthermore, the use of video promotes greater self-reflection than the traditional evaluation and better prepares teachers to play an equal role in the conversation. Although many types of professional development attempt to leverage teacher reflection on practice, they often are prone to the recollection flaw. According to the Center for Education Policy Research (2014),
Reflection on practice relies on the accuracy of memory. Educators must recall the
details of prior lessons to maximize their diagnosis, but those details often fade in
memory. Reflection is best when specific, yet memory can only deliver an adumbrated
version of what happened in any given hour. (p. 7)

One type of practice-based professional development that attempts to mitigate this flaw is the use
of video. Using video frees teachers from having to remember all that happened during the
lesson because they can reexperience the lesson through a video.

**Active learning.** According to Darling-Hammond et al. (2017), effective professional
development often involves active, rather than passive, learning. According to adult learning
theory, active learning makes sense for teachers because it is the kind of learning they encourage
in their own students (Drago-Severson, 2009). Such learning may include strategies such as
think-alouds, reading logs, think-pair-share, jigsaws, text-based student discussion, and problem
solving. Moreover, teachers need tools that allow them to examine their own work (Ball and
Cohen, 1999). The active learning should be highly contextualized and practice based. Such
professional development gives teachers chances to transform their practice rather than layer on
top of old practice. Also, in line with adult learning theory, Hawley and Valli (1999) suggest
that professional development should involve teachers in identifying their learning needs and,
when possible, developing the learning opportunity and even the process to be used. When
video is to be used for professional development, it is often recommended that teachers be
allowed to choose the video that they would like to submit for evaluation (CEPR, 2014; Klein &
Taylor, 2017; Sherin, Linsenmeier, & van Es, 2009). Allowing a teacher this choice increases
their agency in the evaluation process and, therefore, makes them a more active participant.
**Content focused.** Effective professional development is often focused on the content in a discipline or curriculum (Darling-Hammond et al., 2017). Furthermore, it is most effective when situated in teachers’ classrooms when the teachers are with their students and can study student work, test out new curriculum, or study a certain type of pedagogy. The review also found that, ideally, effective PD is aligned with school and district priorities (Darling-Hammond et al., 2017). Similarly, Ball and Cohen (1999) argue that “teaching occurs in particulars—particular students interacting with particular teachers over particular ideas in particular circumstances” (p. 10). It is easy to see how using video for teacher evaluation aligns with this tenet that professional learning must take place in the context of teaching and that teachers need to analyze their own teaching (Ball and Cohen, 1999). In the routine being proposed, the teacher would film the act of themselves teaching a certain type of pedagogy or curriculum in their own classroom and then analyze their own teaching before submitting the video for feedback from an administrator.

**Coaching and expert support.** Darling-Hammond et al. (2017) found that providing coaching and expert support is often an element of effective professional development. This support usually involves a coach or hired consultant sharing expertise about content or teaching practice aimed directly at the teachers’ individual needs. As discussed earlier, the goal of the administrator or supervisor helping teachers develop by providing coaching and expert support through the teacher evaluation process is often subservient to the other goal of teacher evaluation—accountability. The routine that I am proposing would, ideally, help to rebalance those two goals by focusing on the administrator providing contextualized feedback that the teacher would have access to as they rewatch their lesson.
Collaboration. Darling-Hammond et al. (2017) found that effective professional development often involves collaboration in which teachers have space to work together, often in job-embedded contexts. This collaboration might be within a school or across schools, and the group sizes can vary from one-to-one coaching to school-wide collaboration. Collaboration between teachers helps effect changes beyond individual classrooms. Interestingly, Hawley and Valli (1999) found that “although collaborative cultures facilitate school improvement and teacher learning, most schools still isolate teachers from one another most of the time, providing little opportunity for purposeful social interaction” (p. 131). Ball and Cohen (1999) make a similar argument for collaboration. They believe that professional learning should be collective and break down the isolation of teaching through professional discourse that allows for “analysis, criticism, and the communication of ideas, practices, and values” (p. 13). Although video can be used in multiple ways for professional development, many of them (video clubs, mentoring, clinical observations) lend themselves to a collaborative learning experience.

Use of models and modeling. Often in effective professional development, models of effective practice are developed, and then those practices are modeled (Darling-Hammond et al., 2017). Through the use of models and modeling, teachers are provided a clear vision of what best practices look like through either video or written cases of teaching, demonstration lessons, unit or lesson plans, observation of peers, and student work samples. Furthermore, the review found it effective to provide professional learning in conjunction with a model curriculum.

Sustained. Effective professional development is often of sustained duration, providing substantial time for teachers to learn, practice, implement, and then reflect upon new teaching strategies (Darling-Hammond et al., 2017). This is directly opposite to the one-off workshops still prevalent in many schools. The review found that effective programs usually spanned
weeks, months, or even years. The most common model found in the extensive review is an initial, intensive workshop followed by applications in classrooms and additional development days for follow-up.

In summary, the organizational routine that I have designed includes multiple components of effective professional development as defined by Darling-Hammond et al. (2017). First and foremost, it focuses heavily on feedback and reflection by providing a vehicle for observers to give time-stamped, specific feedback to teachers linked to actual teaching while also providing an opportunity for teachers to watch themselves teach and reflect upon that teaching. It also incorporates active learning due to the teacher’s role in the process, is content focused because it focuses on the actual instructional delivery in the context of the classroom, and provides coaching, with the observer being rebranded as a coach providing support.

The Use of Video as a Professional Development Tool

The use of video in teacher education has been around in the United States since the introduction of portable equipment in the 1960s (Sherin & Han, 2004). Over the past two decades, the use of video as a tool both for teacher education and for professional development has grown in popularity (Gaudin & Chaliès, 2015; Hollingsworth & Clarke, 2017; Klein & Taylor, 2017; Muir, Beswick, & Williamson, 2010; Sherin & Han, 2004 Tripp & Rich, 2012a). Video has been used for professional development in many ways. For example, teachers have participated in video clubs in which they collaboratively view clips of classroom practice (Sherin & Han, 2003; Sherin, Linsenmeier, & van Es, 2009; Sherin & van Es, 2009). Another example is creating exemplar teaching videos and making them available to teachers (Grossman, 2013). Video has also been used in various forms in teacher education programs with preservice teachers working alongside mentors and teacher educators (Klein & Taylor, 2017; Muir,
Beswick, & Williamson, 2010; Rhine & Bryant, 2007). Multiple reasons are given for the proliferation of video, including an overall emphasis on practice-based professional development (Sherin & van Es, 2009) and the ability of video to promote skills necessitated by the education reforms of the past 20 years: deeper reflection, better noticing, and improved analysis (Klein & Taylor, 2017; Rhine & Bryant, 2007; Sherin & van Es, 2009). Within this context of using video for professional learning, I will focus on three of the most high-profile uses.

**Best Foot Forward Project.** The Best Foot Forward Project, out of Harvard University’s Center for Education Policy Research, has provided substantial research and some case studies around the use of video for professional learning. This project, which grew out of the Measures of Effective Teaching (MET) study, began by exploring an alternative way to conduct classroom observations (CEPR, 2014). This alternative way, allowing teachers to submit lessons via video for evaluation, then grew into the project, which outlines five ways to leverage video for professional learning: (a) video for self-reflection, (b) video for peer collaboration, (c) video for virtual coaching, (d) video for evaluation, and (e) build a video library. In all of these approaches, the use of video is supposed to break down some of the isolation that has been traditionally found in schools (Lortie, 1975). According to the researchers of the Best Foot Forward Project, this isolation has been a major barrier to improving teaching. They hypothesize that “if used well, video technology can help overcome these barriers and accelerate the process of opening up instruction to observation and feedback” (CEPR, 2014, p. 6).

**National Board Certification.** Although there are limited examples of using video for professional learning, there are more examples of using video as part of a performance assessment leading toward licensure. The most well-known is probably the National Board Certification. 

...
Certification Test, which is over 25 years old. To receive licensure as a National Board Certified teacher, a candidate must submit a portfolio that includes a videotaped lesson and teacher analyses of that lesson. The creators of this assessment, teachers themselves, advocated for the use of a performance assessment rather than a multiple-choice test because they considered the former a better indicator of pedagogical proficiency and more in line with their goal of raising the status of teaching to that of a full profession (Darling Hammond & Hyler, 2013). Similarly, 12 California universities, under a program called the Performance Assessment for California Teachers (PACT), have used performance assessments for licensure since 2002. Like Board certification, PACT required candidates to submit artifacts to demonstrate teaching proficiency, including a videotaped lesson. This work became so popular that it expanded to 30 universities in California, and the idea spread across the nation. Ultimately, a group was formed in collaboration with Stanford University, and the result was a national performance assessment for teacher licensure known as the Educative Teacher Performance Assessment (edTPA) (Darling Hammond & Hyler, 2013).

edTPA. While edTPA has been attacked as part of the “corporatization of teacher education” that is trying to standardize the complex task of teaching, its proponents argue that implementing a rigorous, performance-based assessment system designed by teachers and teacher educators is the best way to lift up the status of teaching as a profession and the quality of teachers entering that profession (Darling Hammond & Hyler, 2013). Although my routine focuses on the development of in-service teachers, I similarly argue that the performance-based nature of videotaped teacher evaluations combined with reflection will both help individual teachers improve their practice and improve the practice of teaching as a whole. Furthermore, I
believe that involving the teachers themselves in the design of the process gives it further
credibility as a routine that may increase the professionalization of teaching in our district.

**Organizational Routines**

As stated earlier, I am using Feldman and Pentland’s (2003) definition of organizational
routine as “a repetitive, recognizable pattern of interdependent actions, involving multiple
actors” (p. 96). The concept of organizational routines provides a helpful way to examine
teaching and the work of teaching. Although historically routines have been thought of as stasis-
producing, stabilizing, and more conducive to rote skill work, in the past two decades routines
have increasingly been seen as a potential catalyst of change and thus appropriate to study in
more complex, highly skilled work (Feldman & Pentland, 2003; Feldman, 2000).

Although teaching is a complex task filled with nonroutine work, organizational routines
can transform teacher practice (Sherer & Spillane, 2011). In fact, although routines are often
thought of as being demotivating, bureaucratic, inflexible, inert, and taking away teacher agency
(Feldman & Pentland, 2003), they can build a culture of professional learning as they become
ingrained into a school culture over time (Sherer & Spillane, 2011). In Sherer and Spillane’s
study, routines also provided constancy and stability over time as they became the scripts that
educators used to discuss instruction and provide a common language.

**Organizational Routine Research in Education**

Several empirical studies explore the role of organizational routines in schools. After
doing a search for research on organizational routines in education, I combed through the list of
results and pulled out the articles that looked at how organizational routines were implemented in
K–12 school environments. Many of the results focused on organizational routines in preschool
or higher education, and I did not include those in my research. I ended up with four studies. In
this section, I synthesize those four studies, paying particular attention to implications that inform the creation of the organizational routine that I have designed. Three of the four studies looked at how changing an organizational routine affects the technical core of a school—instruction (Akiba and Wilkinson, 2016; Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). The fourth (Conley and Enomoto, 2009) looked at the effects of changing a more mundane organizational routine—attendance taking. While not focused on instruction, this study still provided insight into the change process around organizational routines in schools.

The foundational theme I found after reviewing the empirical literature was that organizational structures and routines can indeed be avenues for creating substantive change in schools. Through the design and support of a new organizational routine, an educational leader can change practice (Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). Two of the studies (Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011) found that a change in organizational routine significantly affected instructional practices. The leaders in these two studies designed and introduced organizational routines that included benchmark assessments, grade level meetings, report card review, lesson plan review, morning rounds, review of achievement data, and a revision and implementation of curriculum to ensure vertical and horizontal coherence. The leaders found that these new organizational routines were able to serve as a coupling mechanism that allowed the leaders to modify and leverage governmental regulations to substantially improve instructional delivery in the classroom (Spillane, Parise, & Sherer, 2011). Substantive improvements included standardizing the instructional program, setting and maintaining direction, and monitoring progress by making classroom instruction more transparent. Furthermore, as the organizational routines became institutionalized over time, they provided constancy and stability, because the routines became the scripts that
educators used to discuss instruction and provided a common language for instruction across the school building (Sherer & Spillane, 2011). In some the cases, the routine became so ingrained into the school culture that it was sustained even when school leaders left (Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011).

Akiba and Wilkinson (2016) also focused on the technical core of teaching, instruction, and found that it did not change much when existing organizational routines were kept in place while leaders tried to enact a new, innovative approach to professional development. In their study, Akiba and Wilkinson researched the implementation of Japanese lesson study through the state of Florida. Lesson study, a practice-based professional development approach that is widespread in Japan, involves small teams of teachers developing a common lesson, having one teacher teach the lesson while the others observe, and then conducting a follow-up discussion on the effectiveness of the lesson and how it might be improved moving forward. Akiba and Wilkinson found that implementing this innovative approach to professional development was often difficult, partly because most districts tried to implement lesson study using the existing structures and routines in place and underestimated “the fact that lesson study in Japan is an institutionalized process that is embedded into the organizational structures and routines” (p. 75).

The challenges that the Florida districts faced included teachers’ work schedules, which did not allow adequate time to engage in teacher research such as lesson study; the culture of the teacher as implementer rather than the teacher as researcher; and the lack of support in the form of a more knowledgeable other, whether that person be an instructional coach or another instructional leader. Changes to one or all of these existing routines might have given lesson study a better chance of success.
Although organizational routines have been found to play an important role in changing practice, changes to existing routines are often resisted (Akiba and Wilkinson, 2016; Conley & Enomoto, 2009; Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). Teachers often feel comfortable and familiar with an established routine, and administrators must think through the consequences of altering routines and be prepared for resistance to a new routine (Conley & Enomoto, 2009). In fact, this resistance can jeopardize an organization’s ability to strive for the ideal outcomes of the new routine (Conley & Enomoto, 2009). For example, Conley and Enomoto (2009) examined the effects of a change in a seemingly mundane school routine, attendance taking. They studied a school across the administrations of different school leaders and reported on the effects of changing the way that teachers, secretaries, and administrators managed the daily attendance. Although this topic might not seem controversial, Conley and Enomoto found that the changing of this routine significantly affected the players involved. Another interesting finding was that whereas vaguely defined rules generated more variation in implementation, highly specific rules caused greater resistance to the change. Sherer and Spillane (2011) also reported that at some of the schools they studied, some teachers felt that these new organizational routines undermined the norm of privacy in their instruction that had been respected under the previous administration.

Since a change in organizational routine is often resisted, it would serve administrators well to mitigate implementation issues by being mindful of teacher agency. Administrators would be wise to include staff in the thinking about and reassessing of their existing organizational routines (Conley and Enomoto, 2009). As the people who ultimately perform the routines, teachers can change them. It is therefore important to get feedback from teachers, continually reassess the routine, and address resistance in order for the routine to become
institutionalized in the school. Once institutionalized and part of the school formal structure, these routines can continue to shape practice (Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). Sherer and Spillane (2011) summarize, “If organizational routines are implemented and institutionalized, the values pressed by school leaders through these routines, though more or less challenged at the outset, may become normative over time through the ongoing performance of the routine” (p. 615).

In the next section, I describe an organizational routine that I have designed which is informed by both the lessons learned from the studies outlined above and the tenets of effective professional development described in an earlier section.

**Organizational Routines and My Study**

For my study, I have designed an organizational routine that is based upon my analysis of the reviewed empirical literature, while also incorporating some of the components of effective professional development. The main takeaways from the empirical literature were that organizational routines can bring about substantive and sustained change when enacted with fidelity, that leaders need to anticipate resistance to change, and that leaders should be mindful of teacher agency when designing and implementing a new organizational routine. The organizational routine that I have designed involves using video to reframe teacher evaluation as a form of practice-based professional development. In short, teachers will videotape themselves, choose a lesson on which they would like feedback, complete a self-reflection, share their lesson with an outside observer (me), and then receive feedback directly tied to the video followed by a debrief.

After reviewing the literature, I am more convinced that organizational routines can bring about substantive change and that my particular routine has the opportunity to change teachers’
perceptions of their instructional practice and teachers’ sense of being treated as professionals through the teacher evaluation process. In addition to that foundational and affirming takeaway from the studies, I learned that an educational leader can use organizational routines as a coupling mechanism with outside pressures in order to effect change inside the classroom (Sherer & Spillane, 2011; Spillane, Parise & Sherer, 2011). In my study, I designed an organizational routine to leverage government regulations of the state of New Jersey’s teacher evaluation mandates to increase teacher sense of efficacy of their classroom instructional practice and their sense of being treated as a professional within the context of the teacher evaluation process.

A second big takeaway from the empirical studies was that leaders should be aware that a change in routine is often resisted because most people are comfortable with the status quo. The fact that participation in this study is optional should help mitigate resistance. However, if I attempt to scale up this organizational routine, I will then need to think through how resistance to this change could manifest and possibly derail the intended outcome.

Finally, I have designed my routine so that it involves active learning for the teacher and respects their professional agency. As evidenced by the studies referenced above, teachers play a large role in the performance of any routine affecting instructional practice, and therefore I believe it is important to be mindful and respectful of that teacher agency from the initial design. Therefore, I am framing my study as an opportunity for them to be more actively involved in the design of their own evaluation process and to have a major role in designing the second cycle of the routine.

Conceptual Framework
My conceptual framework, at its core, was very simple. It began with the current state in two areas: teachers’ perception of how the evaluation affected their practice and their sense of being treated as a professional. My working theory was that both teachers’ practice and perception of being treated as a professional could be improved through the introduction of a new organizational routine. Specifically, I designed an organizational routine that uses video for formal teacher observations. Video-based observations allowed flexibility in scheduling and completing the observation cycle. The organizational routine reflected some aspects of effective professional development as outlined by Darling-Hammond et al. (2017), including feedback and reflection, active learning, content focus, and coaching and expert support. During this cycle,
teachers videotaped a lesson after consulting with the administrator (me) on an element of the Stronge teaching framework on which they wanted feedback. The teachers then had opportunity to record as many lessons as they wished and to choose the lesson that they would like to share with the administrator for feedback. This process of choosing their own lesson for observation addressed the need for ownership that is often missing in the feedback cycle (CEPR, 2014). They also received targeted feedback via time-stamped comments that pointed to specific moments in their teaching.

After engaging in this organizational routine, I collected data to better understand whether teachers believed that the quality of feedback they received had improved in such a way that they were better able to improve their practice. What made my conceptual framework slightly more complicated was that, although I designed an intervention in the form a new organizational routine, the teachers themselves almost assuredly did not experience the routine in its idealized (ostensive) form (Sherer and Spillane, 2011). The teachers could influence, but not control, how these routines were enacted in their professional lives. This limitation, referred to as bounded autonomy (Coburn, 2004), affected the implementation of the routine, and thus the performative routine differed from the ostensive routine. Thus, although I designed the ostensive routine, I needed to consider that the performance of an organizational routine could be limited by teacher bounded autonomy (Coburn, 2004). Therefore, the conceptual framework showed the teachers’ current perceptions of the quality of the feedback being filtered through the performative routine, rather than the ostensive routine as teachers experienced a modified version of the routine.
CHAPTER THREE: RESEARCH DESIGN AND METHODS

The purpose of this study was to introduce a new routine into my district, using video for teacher evaluation, and to examine the experiences of the teachers involved and myself, as the administrator who conducted the observations. Therefore, I used a qualitative, action-research design to examine the following research question: “How do we experience the organizational routine that I developed using video for teacher evaluation for improving teacher practice as compared to the existing teacher evaluation system, and how does our experience inform the routine’s continued development?” In the following sections I will describe the research design, including individual decisions about that design and my rationale for those decisions.

Methodological Approach: Qualitative Action Research

A qualitative design was best for this type of research, because I was interested in examining the experiences of the teachers and myself as we participated in this study. Qualitative research allowed me to use both individual and group interviews to collect data in a way that captured the rich, lived experiences of the participants (Merriam & Tisdell, 2016). A key component of this study was the coconstructing with teachers of a routine for the purpose of increasing teachers’ sense of being treated as professionals as well as helping them improve their instructional practice. Due to this component, action research was the best methodological approach to use. Although action research has been defined variously, Herr and Anderson (2015) state that a researcher must state explicitly the definition they are using, because this choice will affect decisions throughout the research process. For the purposes of my dissertation, I used the definition from Argyris and Schon (1991), who frame action research as research that

- is done by practitioners within particular, local practice contexts;
- bounds episodes of research according to the boundaries of the local context; and
• builds descriptions and theories within the practice context itself and tests them there through intervention experiments—that is, through experiments that bear the double burden of testing hypotheses and effecting some (putatively) desired change in the situation.

I chose this definition because it focuses on practitioners working locally to improve a practice, and that is what my participant teachers and I were trying to do.

Action research was the best choice for my study because it aligns well with all three of the circumstances that brought about the practitioner research movement in North America (Herr & Anderson, 2015). These circumstances were the valuation of more systematic qualitative research as a pushback against the dominance of quantitative, positivist research; research showing that collaborative change efforts in the context of schools are more effective than outside initiatives; and a push to “reprofessionalize teaching and to reclaim teachers’ knowledge about practice as valid” (Herr & Anderson, 2015, p. 22). While I studied the organizational routine of teacher evaluation, I also studied the process of developing that new routine.

Therefore, it made sense to frame my study under the umbrella of action research due the fact that I co-developed with teachers a routine for the purpose of increasing teachers’ sense of being treated as professionals as well as helping them improve their instructional practice. Action research allowed me to reflect on both the findings of the research and the process that was undertaken throughout (Herr & Anderson, 2015).

**Context**

The study took place in the Hickory School District (pseudonym), which is located in Hickory, New Jersey. Hickory, a few miles from the New York State border, is located in northwest Bergen County. The Hickory School District is composed of five schools serving
students from grades Pre-K through 12. The overall enrollment of the district is approximately 2,700 students.

**Ethics and Positionality**

I currently serve as the director of curriculum, instruction, and assessment in the Hickory School District. In this capacity, I do not directly supervise any teachers, but I do oversee the entire teacher evaluation process and professional development program. As a person in a position of authority in the school district, I realized it was essential to unpack the role of my positionality before conducting my study. As a participant myself in the action research, I attempted to position myself as an insider/outsider who is aiming for “equitable power relations” (Herr & Anderson, 2015, p. 31). Although I considered myself an insider as a member of the Hickory learning community, I also knew I was viewed as an outsider by the teachers, partly because of my role in the district and partly because I was the one performing the evaluations during the study. Furthermore, I was conscious that during the research process the teachers might only be telling me what they thought that I wanted to hear. Therefore, I thought it best to assume an insider/outsider role while aiming for reciprocal collaboration. I approached my positionality head-on with the participants by clearly stating that their honesty was integral to the process and that, in fact, part of the study would involve them in designing the second (and, postdissertation, third) iteration of this new organizational routine. Thus I hoped to convey to them that there was nothing I necessarily wanted to hear, except for their truth about the process and their experiences. Other ways that I hoped to mitigate any negative effects of my positionality and increase reliability were by triangulating data, by member checking, and by consulting a critical friend.
I triangulated by using multiple methods to collect data, corroborating findings from one source of data by reviewing evidence from other data sources in order to increase the reliability of the data collected (Merriam & Tisdell, 2016). After analyzing the data for each cycle, I shared my meaning making with the participants to see if they agreed with the analysis but also to see how we could use it to inform the next steps in the development process. This process, called member checking, is intended to further the participants’ role as co-researchers in the process (Merriam & Tisdell, 2016). Finally, I consulted a critical friend outside my district’s context to share data and analyses so that they could offer a different perspective and help me see any blind spots that I might be missing.

Participants

I pursued this research with three teachers across the school district who became my co-researchers. For the purpose of anonymity, I asked each of them to choose a pseudonym which is used throughout this dissertation. In May 2019, I asked another district administrator to solicit participation in the study. This administrator solicited participation from tenured teachers who had been rated effective or highly effective in the previous year. From the teachers who volunteered, the administrator chose three who worked in varied teaching positions in the district (grade level, subject, etc.) and only provided me with the names of those teachers; thus, I did not know who had volunteered, and I was able to conduct the research with elementary, middle, and high school teachers representing various academic disciplines. The purpose of limiting the pool to tenured teachers who were rated effective or highly effective was to limit any possible risk that any teacher participant might feel because of being part of the study, given that it was being used as part of their evaluation process. Because we use the Stronge Teacher Evaluation system, the individual observations I conducted did not get a numerical rating. Rather, each teacher
received only narrative feedback tied to specific pedagogical indicators. This is different from other systems, like the Danielson framework, that require the evaluator to rate teachers on specific domains for each observation. Furthermore, the observation I conducted for each teacher was only one of two observations of that teacher for the year, and I provided teachers the option to have a “do-over” if they were not satisfied with the outcome of the video observation. As a final mechanism to protect the participating teachers, I gave them the option to replace their videotaped observation with a traditional in-person observation if they were unhappy with the result. None of the three participants ended up requesting a “do-over” or in-person observation during this study.

After the teachers had been selected, I trained them to use the Swivl and Swivl Cloud platform. A Swivl is a robotic mount that holds an iPhone or iPad that is connected through an application to a device that the teacher wears around their neck. As the teacher moves around the classroom, the Swivl rotates to follow them while the app records video using the camera on the iPhone or iPad. These videos are then saved to the Swivl app and can be uploaded to Swivl Cloud, a cloud-based platform that allows for features such as collaboration and two-way time-stamped commenting. One of the documented issues with using video for professional learning is the difficulty of using video technology in the classroom (CEPR, 2014). However, I used the Swivl and another online platform for an action research pilot study that I did during my doctoral coursework; that training was successful, with the teachers finding it easy to use. I also built in two cycles of this process to the study so that the teachers had one cycle during the spring to practice with the technology before it was officially used for an observation in fall 2019. The teachers did not report any difficulty with the technology.

**Teacher Profiles and Prior Experiences With Existing Teacher Evaluation System**
All the teachers in the study are experienced, tenured teachers who had mixed feelings about the existing teacher evaluation system. Overall, they considered the system a necessary evil. Each could identify both strengths and weaknesses in the system.

**Elise.** At the time of this study, Elise had worked as a fifth-grade teacher for six years in the Hickory School District. She entered the profession later in life, after staying home to raise her family. Prior to her current position, she worked several years as a preschool teacher and a classroom aide. As a humanities teacher, she was teaching reading, writing, and social studies in an interdisciplinary manner at the time of this study. She cited the main influences on her teaching practice as the numerous professional development opportunities in these areas over the prior six years. These opportunities have included participating in summer workshops focused on literacy, writing curriculum for the district, and working with outside consultants who have come in to work with the district. She saw the existing teacher evaluation system as providing feedback on the particular lesson observed but as limited in its ability to contribute to her overall development as a professional. Overall, she viewed the system as lacking ongoing accountability or support and as being transactional and something that must be “checked off” by both the teacher and the observer. In other words, she felt like observations were “one-off” episodes unconnected to the larger picture of professional development. Elise worked with me on a previous action research project involving the use of video for professional learning. In spring 2018, I conducted an action research project involving four teachers working with our district’s literacy consultants and instructional coach on using video for virtual coaching, and Elise referenced this experience as she made sense of her experience with this routine.

**Jonathan.** Jonathan was a high school social studies teacher with more than a decade of experience. At the time of this study he was teaching United States history and a few electives in
the department. He also coached. As a coach he had used video with his team. He noted that in his previous district he never received any constructive feedback after any observation of his teaching, but that in Hickory the norm is that each observation will include at least some mention of areas for growth. A few times during the research process Jonathan referenced how the evaluation system is experienced very differently by nontenured and tenured teachers. Jonathan considered himself an active participant in the current evaluation system because he took seriously the written feedback given to him by supervisors. He believed that this constructive criticism had helped him to improve his practice and thereby earn a higher summative score. Jonathan believed the evaluation system failed to give teachers concrete examples or suggestions for growth. Outside the evaluation process, Jonathan considered the opportunity to collaborate with colleagues in his department as one of the main influences on his teaching practice.

Mark. At the time of the study, Mark was a very experienced middle school English teacher and coach who had used videos to inform his approach to coaching. He cited the main influences on his teaching practice as the trial-and-error experiences of the classroom and in-house professional development. Mark appreciated the current system’s use of professional standards and related performance indicators. He saw the system as informative for a teacher but lacking the ability to help them improve. Overall, he did not see the teacher evaluation system as helping him improve his practice or his sense of being a professional.

The Routines

This action research project consisted of two cycles. The first I designed before the study; the second I left more loosely defined at that time. I did this to allow for the research spiral of the first cycle to play out iteratively so that it could inform the design of the second cycle (Herr & Anderson, 2015, p. 76). I implemented the first cycle with the understanding that
the routine would be changed depending on analysis and feedback from teachers, because the coconstruction of the second iteration of the organizational routine was central to the research. The first cycle was conducted in June 2019, and the second between September and November 2019.

**Initial Version of Routine**

The routine was designed to begin with an individual meeting between myself and each teacher during which we decided on a focus for the first cycle of observations. I allowed each teacher to decide upon their own area of focus by picking from a menu of performance indicators from the Stronge Evaluation System. This served as the preobservation conference part of the formal evaluation process. In our district, teachers must complete a form before this conference. However, for this preobservation process, we decided to forgo the traditional form and collaboratively decided upon a narrower focus on one aspect of instruction. The teachers then had two weeks in June to videotape themselves as many times as they wanted and then choose one lesson to submit to me for review. They also completed a Google Form asking them to reflect on the lesson that they chose to submit. After receiving both the video and the reflection form, I watched the lesson and made comments directly on the online platform we used. Then I contacted the teacher to set up a postobservation debrief. We reviewed parts of the video and discussed comments I had made on it.

**Revised Version of Routine**

The second cycle of the action research process took place during fall 2019. Based on feedback from the teachers and my own reflection on the routine as enacted in the spring, as described in more detail in Chapter 4, we changed the routine for the fall cycle. This feedback and reflection emerged in individual interviews and a final group interview at the end of the first
cycle. Throughout the cycle the teachers were aware that part of the research process was the coconstruction of the revised routine for the second cycle, and therefore they knew that I was interested in their feedback. During the final group interview they suggested improvements for the second cycle, and by the end of the interview we had an agreed-upon routine. It still began with an individual meeting between me and the teacher in which we collaboratively set the instructional focus for the upcoming window during which they would film themselves teaching. Following this initial meeting, the teachers had about five weeks to videotape lessons in their classroom and choose one to submit for their observation. The 5-week window was longer than the spring cycle, because all teachers found the small window in the spring to be a little difficult to navigate. Besides submitting a video through Swivl Cloud, each teacher completed a Google Form containing a few reflective questions. This form was condensed from the first to the second version based upon teacher feedback. After receiving both the video and the reflection form, I watched the lesson and made comments using the time-stamped-comment feature on the online platform. The teachers were notified of these comments at least 24 hours before the scheduled postobservation conference. During this interim they were encouraged to reply to any comments that they felt warranted a reply. We then met, reviewed some of the comments, and collaboratively watched two to three scenes from their submission for further analysis and discussion.

Data Collection

Through both cycles, I utilized multiple forms of data, including a teacher narrative/survey, interviews, observations, and documents (Merriam and Tisdell, 2016). As a reminder, my research question was, How do we experience the organizational routine that I developed using video for teacher evaluation for improving teacher practice as compared to the
existing teacher evaluation system, and how does our experience inform the routine’s continued
development? Throughout the data gathering, I explored both the teachers’ experiences and my
experience with this new organization routine, asking how the teachers experienced the various
components of the new routine, whether teachers could come to see teacher evaluation
differently, and how our collective experiences could inform the development of a second
version.

**Teacher Narrative/Survey**

In late May 2019, I distributed a survey through Google Forms for the teachers to
complete. The questions were open ended and focused on their experiences with the current
teacher evaluation system as it was being implemented in our school district. This survey
informed the initial individual interview.

**Semistructured Interviews**

Following the survey, I conducted a semistructured interview with each teacher that was
audio recorded and then transcribed. The semistructured-interview format is used in qualitative
research that utilizes a mix of more and less structured questions (Merriam & Tisdell, 2016).
The benefit of this format is that it “allows the research to respond to the situation at hand, to the
emerging worldview of the respondent, and to new ideas on the topic” (Merriam & Tisdell, 2016,
p. 111). I used a semistructured interview format because I was addressing some structured,
standardized questions to all the participants, but I needed the flexibility of a semistructured
interview to allow me to follow up on specific things that came up during the interviews.

The interview questions focused on teachers’ perceptions of the efficacy of the teacher
evaluation process currently enacted in the district. For example, questions included the
following: How does the current teacher evaluation system affect your teaching practice in the
classroom? How does the current evaluation system contribute to your development as a teacher? How does the current evaluation system provide feedback or encourage reflection? In what ways do you consider yourself an active participant in the current evaluation system? What are the strengths of the current evaluation system? What do you see as missing from the current system? The interviews were deliberately positioned after the survey in the process so that I could use answers from the surveys to craft specific questions for the participants.

The postobservation conference was immediately followed by another interview. Whereas the postobservation conference focused on participants’ actual teaching during the lesson, the interview focused on their experience with the new routine. Two of the questions that I asked after conducting interviews after the observation were as follows: Are there ways in which you have adjusted your practices in response to the evaluation experience? Can you give examples?

**Researcher-Generated Documents and Artifacts**

A main source of data for this study was various types of documents. Merriam and Tisdell (2016) refer to these as researcher-generated documents, which are “prepared by the researcher or for the researcher by participants after the study [has] begun. The specific purpose for generating documents is to learn more about the situation, person, or event being investigated. They are extremely common in action and participatory research studies” (p. 174). The types of researcher-generated data that were collected during my study were an opening survey, videotaped lessons with time-stamped comments, and accompanying teacher reflections about each lesson.

As a participant in the study, I also kept a reflective researcher journal (Lankshear & Knobel, 2004) in which I documented my thoughts throughout the process so that I had a
memorial of my ideas, including around the development of the second cycle of the routine in collaboration with the teachers. The journal also allowed me to critically reflect throughout and at the end of the process.

**Group Conversations**

The first cycle was completed with a group conversation, conducted and audiotaped during the last week of June. The purpose of this conversation was to get feedback from the teachers about the routine and how it could be improved going into the second cycle. The purpose of having a group conversation instead of another individual interview was to provide participants an opportunity to hear about each other’s experiences and thereby generate new ideas that might not have arisen in a one-on-one setting. Questions focused on their experiences and also asked them to recommend how the routine could be improved for the second cycle. Furthermore, a group conversation designed to provide specific feedback aligned with my goal of coconstructing the second cycle with the participants. Based on the feedback from the group conversation, I refined the routine, which was then used in the second cycle of action research, which we completed in the fall.

**Data Analysis**

As in many action research studies, I analyzed the data as they were collected and then used that analysis to inform the next cycle of the process (Herr & Anderson, 2015). In that way, my analysis was iterative and ongoing (Merriam & Tisdell, 2016). After collecting data during the first cycle using the survey, interview, and videotaped lesson, I entered all the data into Dedoose, a web-based platform for analyzing qualitative research data. I initially analyzed all these data using inductive coding. This process yielded many codes, which I used to guide my questions in the group interview. After analyzing the data the first time using the inductive
codes, I analyzed the data a second time using a few a priori codes, including “prior experiences with evaluation,” “experience with the routine,” and “after the routine” (Saldaña, 2016). A priori coding is used when it is desirable to align data with the conceptual framework, this coding and enables analysis that is directly related to the proposed research questions (Saldaña, 2016).

Using Dedoose, I created documents for each participant for each code. For example, I created separate documents for Jonathan that focused on his prior experiences with evaluation, experience with the routine, and experience after the routine. I did this for each participant. I used these documents to help me formulate follow-up questions to ask during the group interview that ended the first cycle. After conducting this group interview, I again uploaded this transcript to Dedoose and coded it using the same a priori codes mentioned above. I also added a fourth code called “suggestions for changing the routine,” because a significant component of this group interview was revising and coconstructing the routine for the second cycle of the study.

After the second cycle of observations was finished, I again conducted individual interviews and uploaded them into Dedoose. I coded these interviews using the a priori code “after the routine” and also added another code called “experience with second version of the routine.” Differentiating the participants’ experiences between the first and second versions was important because I wanted to document and analyze how the participants experienced each version. From these documents, I did a case report for each participant and organized it by code. When finished with these reports, I conducted a cross-case analysis that allowed me to look for similarities and differences in the participants’ experiences.
CHAPTER FOUR: FINDINGS CHAPTER

This chapter presents the findings from the two cycles of action research enacted over the course of spring and fall 2019. As a reminder, the purpose of this study was to introduce a new routine into my district, using video for teacher evaluation, and to examine the experiences of the teachers involved and myself, as the administrator conducting the observations, as that routine was enacted and revised. The overall research question was, “How do we experience the organizational routine that I developed using video for teacher evaluation for improving teacher practice as compared to the existing teacher evaluation system, and how does our experience inform the routine’s continued development?” The context of the study is my school district, which, while encouraging teacher leadership, is also part of the larger teacher evaluation landscape, which emphasizes accountability over teacher development.

Therefore, two goals of this study were to determine whether this newly constructed routine could increase teachers’ sense of being treated as professionals and whether the routine would also help improve their teacher practice as perceived by themselves. It must also be noted that an important part of this study was the fact that I worked with teachers as co-researchers. Thus, they were active and respected members of the study whose feedback and suggestions were used to design the second version of the routine. Therefore, I was also interested in their experience of being participants in the study. I hoped that this new routine could be expanded for use on the district level in coming years. It was important for me to have teachers in the district participate as collaborators in the development of this routine because I believed they had a real stake in the process and an integral perspective that needed to be considered in the development of a new organizational routine. At the same time, I was interested in documenting my
experiences as a leader involved in this type of collaborative research with teachers as well as my experience as an observer using this new routine within the teacher evaluation paradigm.

I organize my findings into four themes. The first is the existing system: teacher evaluation system as transactional and deprofessionalizing. In the second theme, new routine: cultivating professionalism through teacher agency, I explore how both the first and second versions of the routine increased the participants’ perceptions of being treated as professionals due to increased teacher agency in the teacher evaluation process. The third theme is teacher evaluation as supporting professional development, where I consider how the participants and I experienced the routine in terms of improving teacher practice by aligning with the tenets of effective professional development. In this theme, I attempt to differentiate between effective components that could be attributed to the ostensive and performative routine, respectively. As a reminder, the ostensive routine is the routine in its idealized form, whereas the performative routine is how it is practiced. In the fourth and final theme, trust is central when new evaluation routines are being developed, I examine the relational aspects of the study, specifically how relationships affected our experience and what lessons could be learned from the process.

Before I present these thematic findings, I overview the first version of the routine, the process of collaboratively making changes to the routine, and then the enactment of the second version of the routine. This detailed overview of this process helps to contextualize the subsequent findings.

The Routines

First Version of Routine

As a reminder, the initial routine was designed to begin with an individual meeting between me and each teacher during which we decided on a focus for the first cycle of
observations. Each teacher was allowed to decide upon their own area of focus by picking from a menu of performance indicators from the Stronge Evaluation System. For example, Mark decided to focus on questioning techniques, whereas Jonathan wanted to focus on the teaching of critical thinking. This meeting served as the preobservation-conference part of the formal evaluation process. The teachers then had two weeks in June to videotape themselves as many times as they wanted and then choose one lesson to submit to me for review. They also completed a Google Form asking them to reflect on the lesson that they chose to submit. On receiving the video submission and reflection, I watched the lesson and made comments directly on the online platform we used. Then I contacted the teacher to set up a postobservation debrief. During the postobservation, we reviewed parts of the video and discussed some of the comments made. In this section, I share the experiences of the teachers with the first version of the routine, focusing on recording of the video, choosing the video, completing the reflection form, getting feedback via Swivl Cloud platform, and watching the video collaboratively.

**Recording video.** The teachers found the technical process of recording videos using the Swivl to be teacher friendly. According to Jonathan, “It was simple. It was easy. Just got to make sure to charge. I made sure it was charged the day before that I was going to do lessons” (Jonathan, Interview 2, June 19, 2019). The one minor issue that Jonathan found with the recording process was that, at times, it was difficult to hear what the students were saying if they were not projecting their voices or not located near the teacher. This was the exception rather than the rule, as Jonathan shared: “Other than a couple of cases where the students were really talking lightly in groups, you hear a lot of what the students are saying” (Jonathan, Interview 2, June 19, 2019). During this first cycle Mark also found the recording to go smoothly, except for one of his lessons that he was not able to record because his “iPad ran out of charge at the end of
the day” (Mark, Interview 1, June 6, 2019). Even with this technical glitch, Mark found it fairly easy to resolve. He explained:

There was one hiccup I had. I’ve recorded one class, and it didn’t record my voice for some reason, but that was just me messing with the clicker and I did it wrong. I just didn’t have the clicker on. There were moments where it lost track of me, and it started to spin a little bit. Then it would regain its attention and focus on me. (Mark, Interview 2, June 19, 2019)

Elise reported that the technical part of recording and uploading the video was much easier than during the previous action research project that she participated in. During that project, the teachers were asked to record the video, upload to the online platform called Swivl Cloud, download to their desktop, and then reupload to another platform called TORSH. For this routine we eliminated the latter half of that process; Elise thought that this change made the process much more manageable. Another change that made it easier for Elise was being able to use a dedicated iPad for recording. In this district, each teacher is provided an iPad for professional use. However, Elise found it cumbersome to use that iPad for taping purposes, because she then did not have her own iPad available for other purposes. “To take my iPad in and out of its case, because I use it all the time, was a pain. To have the extra iPad dedicated for this was helpful” (Elise, Interview 2, June 17, 2019). Overall, the technical aspect of recording the video was not an issue for the participants in this study. This an important finding because the participants said that one reason they believed some of their colleagues would hesitate to participate in such a routine was the technology aspect of it.

Completing reflection. Before submitting a video for consideration, the teachers were asked to complete a written reflection. Some examples of questions for the reflection were the
following: What was your area of focus for this lesson? What did you notice in your area of focus while watching your lesson? After viewing the video, what would you consider the next steps for your practice in this area? The full set of reflection questions can be found in the appendix.

Choosing video. After recording their lesson or lessons, the teachers had the opportunity to choose which one they wanted to submit for their observation. This required them to view themselves teaching on camera. This was awkward at first for Elise; she shared, “It’s weird at first. It just is. You see everything, every flaw, but you just have to let that go. You have to let it go and realize the purpose of this to evaluate the teaching” (Elise, Interview 2, June 17, 2019).

In deciding how to choose his video for submission during the spring cycle, Jonathan noted that he watched four or five of the six lessons but that he didn’t watch them all in their entirety except for the lesson that he ultimately decided upon for submission. Instead, he skimmed and focused on some interactions with students. According to Jonathan, there was no reason to pore through the entire lesson on video. In explaining why he did not watch all of the videos in their entirety, he said, “[I] had a pretty good understanding of which one I was going to send you early. I decided pretty quickly on the lesson that I was going to send” (Jonathan, Interview 2, June 19, 2019). Jonathan did not choose the worst or the best lesson for submission, but rather decided to send a video that showcased his students and his pride in them. He expanded upon his thinking in choosing the video for submission:

I was particularly proud of this group. I think it did go very, very well. I didn’t intentionally send you one of the poor ones. You find that when you are sharing information that contains any work of yourself. Of course, your students, you feel like they are a reflection on you, even when they’re not. You definitely don’t want to put that
out into the universe if you feel like you could be showing something better. It was something that I was aware of. My intention, originally, I was going to send you the worst one. Then I watched the videos and I said, “I like this group.” It wasn’t something at the top. It wasn’t something at the bottom. It was some marriage between the two. (Jonathan, Interview 2, June 19, 2019)

Mark took a slightly different approach than both Jonathan and Elise. Although Mark taped himself 12 times, he ended up only watching the one video that he decided upon for submission. This was not his original plan, and he shared, “The hardest part was getting myself to [watch] the video” (Mark, Interview 2, June 20, 2019). He added, “[I] was surprised that I didn’t want to go back and look at myself on video” (Mark, Group Interview 1, June 24, 2019).

I didn’t expect that I’d watch all of them, but after having filmed for 11 lessons, I felt like I had a pretty good sense of the rhythm of each one. I could have chosen any one. I would have been comfortable with anyone. (Mark, Interview 2, June 20, 2019)

Even though he did not watch all 11 lessons, Mark said, “[I] spent much more time polishing it than I would have [if] it was an ordinary administrative observation” (Mark, Interview 2, June 20, 2019).

**Getting feedback.** When reflecting upon the existing routine, Elise said that the written feedback aligned to professional standards was one of the few positives. For the first version of the new routine, she would still be getting written feedback aligned to standards, but now it would be tied to a moment in the video. In the existing teacher evaluation system this is possible but very cumbersome to do.

**Watching video collaboratively.** The final part of the routine was the postobservation conference. In the traditional postobservation conference, the teacher and supervisor might have
a conversation about the lesson in question. This postobservation is supposed to take place within five days of the actual observations. Sometimes the supervisor will share their observation report with the teacher in advance of the postobservation, but other times the teacher does not see the report until after the postobservation. For our routine the report was shared ahead of time with all the comments. Then during the postobservation, the teacher and I watched the video together, focusing on the comments included by me ahead of time.

Changes to Routine

A key component of this study was the fact that the second version of the routine would be coconstructed with the teachers after they had participated in the first version of the routine. At the end of the first cycle in the spring, the participants and I critiqued the version that I designed to make modifications and adjustments. In this way we coconstructed the second version, which was enacted during the fall cycle, during September and October 2019. Although the teachers and I had overwhelmingly positive experiences during the first version of the routine, we were able to coconstruct a second version that we believed would even be more professionalizing for the teachers and conducive to improving teacher practice as perceived by the teachers themselves.

Table 4.1

Comparison of First and Second Versions of Routine

<table>
<thead>
<tr>
<th>First Version</th>
<th>Second Version</th>
<th>Rationale for Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first step was meeting to set the focus collaboratively. The focus should be limited in scope and be standards-</td>
<td>The first step was meeting to set the focus collaboratively. The focus should be limited in scope and be standards-</td>
<td>No change.</td>
</tr>
<tr>
<td>Based aligned with the Stronge indicators.</td>
<td>Based aligned with the Stronge indicators.</td>
<td>We extended the window based on teacher feedback from first cycle. This allowed for even great teacher flexibility and control over the process.</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>The teachers videotaped as many lessons as they want over the course of 2 weeks. There was no minimum or maximum number of lessons to tape.</td>
<td>The teachers videotaped as many lessons as they want over the course of 5 weeks. There was no minimum or maximum number of lessons to tape.</td>
<td>No change.</td>
</tr>
<tr>
<td>The teachers chose one 20-minute clip to share via Swivl Cloud. They were to keep the clip small (between 20 and 30 minutes). It had to be at least 20 minutes to satisfy AchieveNJ requirements.</td>
<td>The teachers chose one 20-minute clip to share via Swivl Cloud. They were to keep the clip small (between 20 and 30 minutes). It had to be at least 20 minutes to satisfy AchieveNJ requirements.</td>
<td>No change.</td>
</tr>
<tr>
<td>The teacher completed a reflection via Google Form.</td>
<td>The teacher completed a reflection via Google Form.</td>
<td>The reflection was made more concise based on teacher feedback from first cycle. The process of reflection was considered valuable but redundant, so questions were eliminated to streamline process.</td>
</tr>
<tr>
<td>Teachers received comments.</td>
<td>At least 24 hours before postobservation conference the teacher received time-stamped comments via Swivl Cloud.</td>
<td>The 24 hour notice was based on teacher feedback. The teachers appreciated having time to review and digest comments before coming into the postobservation conference to review and discuss.</td>
</tr>
<tr>
<td>Once the teacher reviewed the comments, they had the option to respond to comments directly in Swivl Cloud.</td>
<td></td>
<td>This new addition was designed to make the observation more of a conversation between teacher and observation. It gave the teacher the opportunity to provide clarification and/or more information when needed.</td>
</tr>
</tbody>
</table>
Had postobservation conference

Had postobservation conference that was focused on a few scenes in the lesson.

Focusing on fewer scenes would allow for more in-depth conversation and analysis.

Teacher submitted another reflection, submitted another video, or submitted artifact to show responsiveness to comments/feedback. (optional)

This option would allow teachers to show evidence of growth and could help observation be part of larger growth conversation.

---

**Second Version of Routine**

The second version of routine, though similar to the first, incorporated five specific changes based on the teachers’ feedback after the first cycle: (a) a longer window of time for the teachers to film their lesson(s), (b) an awareness of and emphasis on teachers using the comment feature in Swivl Cloud to respond to comments, (c) a more concise reflection, (d) a focus on fewer scenes during the collaborative viewing during the postobservation, and (e) an optional reflection or submission of artifact following the observation cycle. In this section, I explain the rationale for those changes and how those changes were then experienced by the teachers and me during the second cycle of action research.

**More time.** One of the changes to the routine was increasing the length of the window that teachers had to complete the process. Although all three teachers thought this was a good idea, Mark especially championed this change due to the time constraints he felt during the first version. We briefly discussed the idea of determining a minimum number of videos that a teacher should video and watch during the time frame, but the teachers did not think that would be helpful.
**Responding to comments within Swivl Cloud.** During the spring cycle, we realized that one of the interesting features of the Swivl Cloud platform is the opportunity for the teacher to respond directly to individual comments and questions, making the evaluation more of a dialogue between teacher and administrators; teachers were encouraged to take advantage of this option during the fall cycle. While the teachers had the opportunity to respond to comments during the first version of the cycle, this was not a feature of the platform that was discussed or emphasized. As we collaborated on the process during the first version, it became apparent to both teachers and me that this was a powerful tool that should be utilized better going forward.

**Making reflection more concise.** During the first version of the routine, the teachers were asked to complete a reflection before submitting their video. They found this reflection process helped them choose which video they wanted to send and allowed them to be more focused in their viewing than they might have been without it. They recommended keeping it for the second version of the routine with some small changes. Although it was not a big issue for any of the participants, we did make the reflection survey more concise for the second version by removing a few questions that seemed redundant.

**Focusing on fewer scenes during collaborative viewing.** During the first version of the routine, I made about 20 comments on each individual teacher’s lesson. During our collaborative viewing, we discussed each of these comments, stopping at each point in the lesson and watching the action to which the comment was attached. During the follow-up interview after the postobservation, Jonathan suggested that, going forward, there be a focus on fewer video clips during the postobservation conference. He made the case that this would allow for a deeper analysis of and conversation about those moments that were discussed. Through the comments they received beforehand, he argued:
The teacher understands what you’re looking for. The teacher will perceive you as saying, ‘Okay. This is what I want you to continue, and then, of course, the areas of improvement with certain examples throughout the lesson.’ I don’t think you need to watch as many. (Jonathan, Interview 2, June 19, 2019)

He thought the number of comments was appropriate and helpful, but that it was not necessary to discuss all of them in the same detail during the postconference. For the second version of the routine, he suggested focusing on two to three points in the lesson to watch collaboratively rather than reviewing all the comments made throughout the lesson. As a coach, Jonathan related this back to his own experiences in using video effectively:

I don’t show the kids the whole game. I show them the 15 clips that I want to show them. Some are good. Some are where we need areas of improvement. That’s exactly what you’ve done here. You’ve set it up where I see the good stuff that I want to keep doing. Then I see the areas where you obviously pause, and it makes you think a little bit. I also see where you had particular questions. It gave me the opportunity to answer those specific questions. There were areas where we stopped, and there were areas for growth that were also built in. It’s clear as day when those areas show up on film.

(Jonathan, Interview 2, June 19, 2019)

This was a change that we did incorporate into the second version of the routine along with a few other teacher recommendations. When this suggestion was raised during the group interview, neither Elise nor Mark had perceived this as an issue, but neither disagreed with the idea that a focus on fewer clips could be beneficial.

**Optional reflection or submission of artifact.** To make the evaluation routine even more collaborative and to increase teacher agency in the process, we discussed the option of
teachers being able to submit a reflection or artifact after the postconference if they felt it was warranted. Now that I have reviewed the components of the routines, I move to my first theme of how the existing teacher evaluation is transactional and deprofessionalizing.

**Existing System: Teacher Evaluation as Transactional and Deprofessionalizing**

As discussed in chapter 2, teaching in the United States has been organized as a bureaucracy, with the power located at the top of a hierarchical structure with high levels of managerial control (Mehta, 2013). Furthermore, this has been coupled with a dominant conception of teacher learning that focuses on knowledge-for-practice, which is characterized by experts outside the classroom transmitting knowledge to teachers and leaders then expecting teachers to implement this knowledge in the classroom (Cochran-Smith & Lytle, 1999). It is not surprising that what has followed is a teacher evaluation system that locates the power in the managerial level and does not treat teachers as professionals by giving them significant control over decisions as part of the process. Furthermore, our current era of oversight, accountability, and testing has led to a teacher evaluation system that skews toward the accountability end of the spectrum and further deprofessionalizes teachers (Donaldson and Donaldson, 2012; Woulfin et al., 2016). The teachers in my study described an existing teacher evaluation system that did not respect their expertise or position them as knowledgeable partners in the process. This is reflected in my first theme: *teacher evaluation as transactional and deprofessionalizing.*

The participants felt that teacher evaluation was something “done” to them and a process in which they were not active participants. This is more in line with how evaluations take place in contexts of low-level, routine work, rather than of the highly skilled work that teaching requires (Mehta, 2013). In this context, the participants shared that they found the existing system to be very transactional and lacking any ongoing or systematic approach. The teachers
also viewed the evaluations as something that had to be checked off for both the administrator and the teacher, with the understanding that it was a waste of time and just something that had to be done as a bureaucratic exercise to satisfy the state requirement. Furthermore, the teachers did not believe that the existing teacher evaluation system contributed to their professional growth. However, despite this overall negative attitude toward the existing evaluation system and its relation to professionalism, the participants did find some value in it, such as providing a common language through shared teacher standards and having an organized structure with clear expectations for teacher performance.

**Passive Recipients: Evaluations Are Something “Done” to You**

Participants described the existing system as one that treated them as passive recipients and that was defined by managerial oversight rather than treating them as valued professionals with knowledge and expertise. When asked, one of the participants, Elise, was “not sure the current evaluation system actually [did] contribute to [her] sense of being a professional” (Elise, Survey, June 2, 2019). Therefore, she did not consider herself a very active participant in the process. For Elise, the current observation system was something done to do her rather than something she was a part of, because she did not see herself participating in the process other than submitting paperwork, teaching the lesson, and afterward receiving the observation report with comments from the observer. This finding is consistent with other research that suggests evaluations are often something that is “done” to teachers (Papay, 2012; Wise et al., 1984). Elise was communicating frustration with the existing evaluation system by saying that it did not encourage her to take an active role in her own evaluation and that it did not respect her own expertise.
According to Jonathan, the current evaluation system was not seen as an opportunity for growth by many of his colleagues at the high school. He attributed this partly to the fact that each teacher receives a numerical rating of their performance at the end of each year and that this numerical rating adversely affects the rest of the system:

The problem with the evaluation system is that people . . . want the number and they want to make sure they’re in good standing. That’s all they really want to know. Again, the problem with the observation system is that they see these indicators, and they’re trying to make the connection between these indicators and what they’re seeing in their lessons. You’re not seeing too many conversations with teachers in the building, talking about where they can grow. Instead it’s just the outcome. (Jonathan, Interview 2, June 19, 2019)

Jonathan here alludes to the fact that the existing evaluation system culminates in a numerical score given to the teacher by their administrator at the end of each year. Along with this numerical rating comes a label of highly effective, effective, partially effective, or ineffective. This structure has come to dominate how the system is implemented and perceived by the teachers. With the looming score coming at the end of each year, the observations lose their ability to serve as a formative tool.

A Bureaucratic Exercise: “Something to Be Checked Off”

The participants shared that they considered the existing teacher evaluation system to be mostly a bureaucratic exercise that had to be done to satisfy the state mandate. It was simply more paperwork and a hoop that needed to be jumped through by both the administrator and the teacher involved. As such, it was seen as an overall deprofessionalizing process that neither
Mark believed that the existing teacher evaluation had devolved into something that held minimal value for teachers and administrators, simply an exercise in compliance, and that it lacked active participation from the teacher. He shared how the process of evaluation sometimes occurred:

Sometimes when the administrator goes into your room, at first, they’re looking at the overall sequence of your lesson like, “OK. Now they’ve gone through all the steps,” and they’re checking off. They’re not listening to the way you’re presenting and what you’re saying. (Mark, Interview 2, June 20, 2019)

Mark was speaking to a system focused on completion and compliance rather than on a process that treated teachers as professionals with the knowledge, skills, and expertise to play an active role in their own evaluation. He did not see connections being made between his observations and his other avenues of professional development, such as the coaching he was participating in or the professional development plan (PDP) that he developed collaboratively with his supervisor.

Elise saw the teacher evaluation system in place at the time of this study as providing feedback on the particular lesson observed but limited in its ability to contribute to her overall development as a professional. Overall, she viewed the existing system as lacking ongoing support and as being transactional and something that must be “checked off” by both the teacher and observer. When asked what was missing from the current process, her answer was:
Firstly, ongoing accountability. I’m not saying the solution is to put cameras in the room. I guess what I mean by that is that once the observations are over, there is sort of a mental check mark: ‘That’s done. Move on’. (Elise, Survey, June 2, 2019).

In other words, she felt like observations were “one-off” episodes that were not connected to the larger picture of professional development and for which there was rarely follow-up. She also shared that another negative of the current system was the time-consuming paperwork involved:

That is very time-consuming. I mean, it is for me. . . . To say, “Okay, so we’re going to increase your observations from 2 a year to 10 a year”—I might just lay down and die. It’s just a lot of time. Then all you’re doing is increasing it 10 instances of 40 minutes. I don’t know that that’s really . . . I don’t know what the answer is to make sure there’s more accountability, but also not more paperwork. (Elise, Interview 1, June 5, 2019)

Although Elise saw ongoing support as missing, she was not sure of a solution, although it was clear that the answer for her was not more observations.

**Trying to Satisfy Competing Purposes: “Something Else Would Have to Be Done”**

The participants described the existing system as one that was not productive for their professional growth. Elise provided a very insightful answer when asked if she felt the existing teacher evaluation system was effective:

I guess that would depend on, what’s the purpose of the observation. Because if the purpose of the observation is to show up and make sure that I’m following the pace and that I’m getting everything done on the lesson plan, then the current system is working. If the observations are meant to do something else, then something else would have to be done. (Elise, Interview 1, June 5, 2019)
Thus Elise articulated that in order to judge the efficacy of the teacher evaluation system, one must first identify the purpose of the system. As we explored in earlier chapters, two competing purposes of teacher evaluation have been accountability and teacher development, with the pendulum swinging back and forth throughout the decades (Donaldson and Donaldson, 2012; Woulfin et al., 2016). In her answer, it can be inferred that Elise believed that the existing system succeeded in fulfilling the accountability purpose but failed to encourage development. We will return to this in the second theme—linking teacher evaluation and professional development.

An aspect of the current evaluation system that Jonathan found deprofessionalizing was the unannounced observation. These observations are required by TEACHNJ (NJDOE, 2014a). He believed that the unannounced observation was unfair to teachers and not productive to teacher growth. During our interview, he recounted a story of a colleague who approached him after I had done a walkthrough of his classroom one morning:

He came up upset as can be. Not because he felt like he was doing anything wrong. Not because he felt like his class wasn’t engaged. Not because he felt like he was teaching a bad lesson. It just didn’t represent what he was as a teacher. That’s what you get unfortunately sometimes when you do these unannounced observations. You can catch a teacher on an odd day. You can catch a teacher who’s maybe—and I hate to say these terms—not at their best. Sometimes you don’t take into account (a) why the teacher is teaching the lesson and why they’re teaching it the way that they’re teaching it and (b) they’re not sure what you’re looking for. That’s why I think that the preconference and the long observation and the announced observation is way more valuable. (Jonathan, Interview 1, June 6, 2019)
This quote illustrates how the teachers view the unannounced observations as drive-bys by administrators who will use a short visit to generalize on a teacher’s overall body of work. His comments suggested that this form of observation served to further deprofessionalize the process by creating an environment in which teachers live in a state of anxiety that an administrator might walk into their class at any minute for a formal observation. To him, this took away all control from the teacher in the process of their own evaluation and growth. It is no wonder that Jonathan’s perception was that the existing teacher evaluation system did not foster an appropriate sense of professionalism.

**Moving Toward Professionalism**

While not finding the existing system strongly encouraging to professionalism, Mark considered the system more so than the previous system in that the existing system was based on teaching standards and transparent indicators. For Mark, having agreed-upon standards made the evaluation process more professional: “The domains give me areas to target and develop. . . . We look at our classroom practices and teaching through the lens of the domains” (Mark, Survey, May 29, 2019). When asked to remember a time before the use of a standards-based system, Mark laughed and implied that those were the dark ages when teacher evaluation was much less systematic and less professional than the existing system. The existing evaluation system made him feel more professional because of the transparency the indicators afforded him in comparison to the lack of a common language of the previous system. This common language was one of the characteristics of the standards-based teacher evaluation system that derived from the work of Hunter and Danielson and came to dominate teacher evaluation systems across the country over the last decade (Stronge and Associates, 2018). While these standards-based systems have been implemented in a myriad of ways, the allure of them is their ability to provide
a common language and framework that promotes both a more professional teacher evaluation process and a better way to improve practice (Donaldson, 2016; Papay, 2012; Toch and Rothman, 2008).

From Jonathan’s perspective another strength of the current evaluation system was the structures it provided for both teachers and administrators. He believed that the organized structure with clear expectations for teacher growth contributed to the professionalization of teaching:

It’s important that people continuously grow in their profession, no matter what their profession is. The evaluation system makes sure that people are doing so. Educators are not the only people that are asked to do PDPs. Educators are not the only people who are asked to show their development over time. (Jonathan, Interview 1, June 6, 2019)

As discussed in the literature review, teacher evaluation has served two purposes — teacher development and accountability (Donaldson and Donaldson, 2012; Marzano et al., 2011; Woulfin et al., 2016). Jonathan’s comments suggested that he was aware of the accountability piece of the existing evaluation system and believed that this accountability was one of the pieces that make the system feel more professional. He saw in this accountability an opportunity for growth. Whereas this one piece of the existing system was seen as a move toward professionalizing, my second theme focuses on how multiple aspects the new routines promote professionalism through the insertion of teacher agency into the teacher evaluation system.

**New Routine: Cultivating Professionalism Through Teacher Agency**

One of the main purposes of this study was to create a routine that increased the professionalism of the teacher evaluation process. For the purposes of this study, I define a
profession as “characterized by deep levels of knowledge and expertise, professionally shared standards of appropriate practice, and the use of judgment and discretion in applying that common knowledge to particular situations” (Mehta, 2013, p. 468). One of the major findings of my study was that the new routine that we coconstructed led the teachers to describe feeling a greater sense of professionalism relative to their experience of the existing teacher evaluation system. The teachers attributed this greater sense of professionalism in part to the increased teacher agency that this routine fostered. I identified three subthemes related to this theme: (a) increased teacher agency: “when you have the choice you have a lot of the power”; (b) improving their own practice through self-critique and reflection: “I’ve taught myself”; and (c) ongoing practice: “preparing yourself as any other professional would.”

Although the teachers in this study had a mixed view of the existing evaluation system and its role in their treatment as a professional, all three teachers in this study were unequivocal that their experience with the research project and new evaluation routine increased their sense of being treated as a professional. Although all three described an increased sense of professionalism, Jonathan put it most succinctly: “To me, this couldn’t haven’t worked out better from a professional standpoint” (Jonathan, Interview 3, October 22, 2019). They attributed this increased sense of professionalism to several factors, including increased teacher agency as a result of increased control over the process, respect of their expertise through affording them the opportunity to critique themselves through reflection, and the chance to receive targeted feedback and recommendations and then practice in a professional setting the new patterns that they committed to as a result of the feedback and recommendations.

**Increased Teacher Agency: “When You Have the Choice, You Have a Lot of Power”**
All three participants commented on how they appreciated having more agency or control during the process. An increase in teacher agency has been shown to contribute to feelings of being treated as a professional during the teacher evaluation process (Biesta & Tedder, 2007). In designing the ostensive routine, which is the routine as it is designed in its ideal form, I purposefully built in multiple ways to increase teacher choice, including choice over how many lessons to record, which lessons to record, and which lesson to share. As the routine evolved, even more choice was built in, in response to feedback from participants.

One of the most important aspects of the ostensive routine was the teachers’ ability to choose which video they submitted for review and the extended time during which they could record videos. As noted earlier, during this first version of the routine, the teachers had two weeks to record as many videos as they wished and then were asked to choose one that they wanted to submit. This is different from the existing teacher evaluation system, where a teacher has either little or no control over when they will be formally observed as part of the evaluation process.

Throughout the spring cycle, Elise recorded six lessons during the two-week window in June. She watched each video and selected her best for submission. Elise really appreciated the ability to choose her video and expressed that “having different things to choose from allows you to choose what you think is the best option” (Elise, Interview 2, June 17, 2019). When asked why she chose to submit her best video, she explained, “Human nature. You want to submit the best one because you want to impress the best. You want whoever’s evaluating you to see what you feel your best effort looks like” (Elise, Interview 2, June 17, 2019). When asked to justify why she predicted she would again choose her best lesson for submission the fall cycle, Elise explained, “I want you to see what I’m really capable of, what my students are capable of doing
because we work together. Yeah, I want you to see the best” (Elise, Group Interview 1, June 24, 2019). For Elise, she wanted to put her best foot forward and viewed the observation as an opportunity to impress the observer with her work.

The participants noted the importance of the changes made in the second routine to increase teacher agency. At our final interview, Jonathan shared, “When you have the choice, you have a lot of power. You feel a lot better about the whole process” (Jonathan, Group Interview 2, October 30, 2019). This sentiment was echoed by Mark: “We had a choice and we had control. That’s why we were more trusting of the process” (Mark, Group Interview 2, October 30, 2019). This issue of control was one that the participants returned to multiple times over multiple interviews. For them, the fact that they were given control over many of the decisions in the process substantially affected how they viewed the routine in terms of whether they felt treated as a professional. This fits with Mehta’s (2013) comparison of bureaucracy and profession and how in a true profession, the professionals have a great deal of control over their own evaluation. This has not been true in many forms of teacher evaluation.

As described in an earlier section, each teacher used this structure slightly differently. Regardless of their specific approach, each teacher ended up taping more than one lesson per day. Mark found this pace a little taxing and observed that it significantly increased the time he needed to prepare for the lessons. He also noted:

Early on when I was recording the students, they were very muted in terms of their responses. They were very aware they were being recorded. They didn’t want to say anything. They didn’t want to interrupt anything. By Thursday and the 12th taping, it was more comfortable. (Mark, Interview 2, June 20, 2019)
We briefly discussed the idea of determining a minimum number of videos that a teacher should record and watch during the time frame, but the teachers did not think that would be helpful. They felt that it would reduce teacher agency. Jonathan spoke to this idea: “Going back to the idea of professionalism through this, the teacher should have the freedom to watch what they choose and don’t watch what they don’t feel like” (Jonathan, Group Interview 1, June 24, 2019). Furthermore, the teachers felt that teachers would end up filming and watching more “when you don’t ask” (Mark, Group Interview 1, June 24, 2019). Jonathan agreed, adding that without having a minimum number assigned, “we just did more. I recorded classes that I didn’t intend to, just to see how it went” (Jonathan, Group Interview 1, June 24, 2019). Based on that feedback, we did not set a minimum number for the second version of the routine.

Along with having control over the number of lessons they filmed, the teachers had control over when they filmed the lessons. In the second version of the routine, the teachers had a five-week window, which gave them increased flexibility in how they could approach the filming. With this increased flexibility, Jonathan explained, “[I] figured out what lesson I wanted to be filmed and waited for my courses to get to those lessons naturally” (Jonathan, Interview 3, October 22, 2019). He found this to be a real positive, because:

sometimes when you have an observation you force a lesson too early or you’re waiting for it because you know an administrator is coming in. . . . Knowing that I had the time frame to work with, I naturally got my students prepared for the lessons that we were going to be recording. (Jonathan, Group Interview 2, October 30, 2019)

Whereas Elise watched each lesson after she filmed it, Jonathan filmed all six lessons before sitting down to watch them all, which he did in one sitting. And whereas Mark did not
watch his lesson until a month after filming it, Jonathan filmed his lessons during two days and then watched them the following day.

One of the changes to the routine was increasing the length of the time window that teachers had to complete the process. Although all three teachers thought this was a good idea, Mark especially championed this change due to the time constraints he felt during the first version. In the individual interview following the first cycle, Mark candidly shared:

I didn’t feel like watching video of myself after a week of recording. I needed some separation. However, as time goes on, I will become more interested. I know you’re under a time constraint to review this, but I feel time away from these recordings will add value to my reflection. (Mark, Interview 2, June 20, 2019)

Although we did decide to increase the time the teachers had to complete their part of the evaluation process, we did so mostly in order to give them flexibility in when and how often they wanted to film. I had not expected someone to take the approach that Mark did of taping a lesson and then not returning to view it for a few weeks. However, this flexibility in time allowed him to do just that, and he found that space between the teaching and viewing to be helpful for him, because it allowed him to more effectively reflect upon his practice. He found that coming back to the lesson after a hiatus allowed him to be more reflective than he would have been if he watched the lesson in the day or two following the actual taped lesson.

All three teachers found the extended window of the second version to be preferable to the shorter window of the first version. Jonathan put it well:

It was great. I had the time frame. It was a relaxing process. It wasn’t stressful. That’s why I like it. When it comes down to the timeframe and also the filming itself, it’s very teacher friendly. There’s no real stressful component to it. You can control the narrative
and the timeline in your own way. It made it easy. (Jonathan, Group Interview 2, October 30, 2019)

Here, again, Jonathan speaks to the importance of teacher agency in the evaluation process. Without increased teacher agency with control over substantial decisions in the process, a teacher evaluation system will continue to be deprofessionalizing.

With the goal of making the evaluation routine even more collaborative and to increase teacher agency in the process, as described earlier, we discussed the option of teachers being able to submit a reflection or artifact after the postconference if they believed that that action was warranted. Elise thought this would be a positive change, because it would allow the teacher to show growth in an area that the observer focused on during the observation. She believed that having the opportunity to submit something “brings it full circle” (Elise, Group Interview 1, June 24, 2019). Jonathan agreed but thought it was very important that this remain optional and not required: “You want to bring it full circle while also at the same time not putting more things on a teacher’s plate. That’s a thing that they might be hesitant on” (Jonathan, Group Interview 1, June 24, 2019). We decided to include this optional reflection or submission of artifact during the second version of the routine. However, although all the teachers thought it was a good idea to include an optional reflection, none of them actually submitted one.

The increased choice and control changed observations for evaluation purposes from something that was being done to the teachers to something they felt a part of. In an upcoming section, we will explore the teachers’ experience being part of the research study and, as a result, a more active participant in their own evaluation. Sometimes the teachers approached the process differently than I had expected, and in these situations I became willing to give up control.

**Improving Teacher Practice Through Self-Critique and Reflection: “I’ve Taught Myself”**
The second aspect of the routine that the participants noted as a factor in increasing their sense of professionalism was that the process encouraged them to reflect upon and critique themselves. For me this was one of the most powerful parts of the routine, because it truly made the teachers active participants in their own evaluations. By the time they ended up submitting their videos to me, they had already reflected extensively through viewing, analyzing, and self-critique. I also felt that the teachers were more accepting of external feedback when it did come. Although the routine did include a written reflection for teachers to complete before submitting their video, this component of reflection and self-critique emerged as something that the teachers found themselves engaged in far more than the routine demanded of them. As discussed earlier, the teachers were asked to film as many lessons as they wanted to over the observation window and then choose one lesson of about 20 minutes to submit as their official observation along with the written reflection. Besides those requirements, the teachers had flexibility in how they wanted to go about the work.

All three teachers approached the task differently, and in a way that made the most sense for them. Elise approached it the way that I anticipated all of the teachers would. She taped herself teaching several similar lessons with small groups of students. She watched herself after each lesson and tweaked her practice based on her observations. Going through this iterative process allowed her to change her practice before even submitting a video. She explained how this process helped her:

If I hadn’t done that over a period of time, I don’t know if I would have been so likely to revise and edit a few times. I definitely got better over time. Doing it a few times helped me practice, and it helped me identify areas to improve. (Elise, Interview 2, June 17, 2019)
She then viewed the video she submitted as the culminating artifact that she wanted to share with me, the observer, as evidence of all that she had learned:

I like the experience that I had. The growth that I showed as I did [the lesson] the first time was like, “Wait, let’s do that differently this time.” It allowed me to make improvements almost to teach myself and then let you see what I’ve taught myself. You see the best part of me and the growth that I’ve had for this particular event that I filmed. (Elise, Group Interview 2, October 30, 2019)

For Elise, the iterative process she followed helped her refine her practice and ultimately gave her confidence that she was improving and that her improvements would stick. She became a student of her own practice, as described by Ball and Cohen (1999). This worked for Elise, who came to see this as a process during which she taught herself how to improve her practice and then viewed the video she submitted as the culminating artifact that she wanted to share with me, the observer, as evidence of all that she had learned during the process:

If what you really want to accomplish is to teacher practice improve, I would say that it does. It does because I watch myself. . . . I did this a few times. I recorded myself doing this a few times and it seems like it creates an accountability to say, “If I wanted my best to be seen, I think just like anything else, you got to practice.” So doing it a few times helped me practice, and it helped me identify areas to improve . . . this helps me feel more confident that the improvements I’m making are effective. (Elise, Interview 3, October 25, 2019)

Elise was a more active participant in her own evaluation and through that active participation was able to teach herself even before submitting the lesson to be viewed by the administrator. In this way, her professional expertise was both respected and utilized as a key component of her
own evaluation and professional growth. This emphasis on expertise put her experience more in line with what teacher evaluation would look like if teaching was regarded as a profession, rather than a bureaucracy (Mehta, 2013). As a professional with a certain level of knowledge and expertise, Elise should have control over her evaluation and development.

Jonathan agreed that the routine ended up being powerful in no small part because the teachers were motivated to watch themselves on video and critique themselves before submitting a video. Unlike Elise, Jonathan filmed all of his lessons over a two-day period and did not view them until they were all finished. He then set aside some time to view the lessons with a critical eye and determine which lesson he wanted to submit. Due to this process, he found that most of his own growth happened “on [his] own. By the time we got to the post-observation there was little to actually say about the topic. We had done a lot of the growth on our own through the use of video” (Jonathan, Group Interview 2, October 30, 2019).

The promise of using video to promote reflection and better analysis skills of teachers when they’re viewing their own practice is something that has been documented in multiple studies (Klein & Taylor, 2017; Rhine & Bryant, 2007; Sherin & van Es, 2009). Jonathan believed that completing the reflection questions allowed the postobservation to be more conversational because he was contributing to his own evaluation by completing the reflection before even receiving comments from the supervisor. “It becomes more conversational, and it becomes more beneficial and helpful” (Jonathan, Interview 2, June 19, 2019). Elise considered the process of self-critique one of the most important elements that the routine promoted:

It lends itself to your sense of being a professional because, in a sense, you’re critiquing yourself. You’re thinking of, “What can I do to make this better?” That’s our job.

Reflecting on . . . the video and looking at it through that lens as if I’m observing myself
makes me feel my thoughts on that are important. (Elise, Group Interview 2, October 30, 2019)

Perhaps due to this reframing of the evaluation paradigm, although the teachers were asked to film as many lessons as they wanted to and then submit one for their actual observation, the teachers ended up putting much more time and energy into the process than they would for a typical observation in the existing system.

**Ongoing Practice: “Preparing Yourself as Any Other Professional Would”**

The third major way that the routine contributed to an increased sense of professionalism was that it lent itself to a more ongoing process, rather than the one-and-done observation that is typical in the existing teacher evaluation. This result aligns with the findings of Darling-Hammond et al. (2017) that effective professional development is sustained and takes place over a period of time. Although no follow-up was required, the routine fostered this. At the final group interview, Mark shared:

> You’re preparing yourself as any other professional would—a lawyer, a doctor. It’s like, “Okay, I’m practicing these nonverbal cues.” You got to read up a little bit. You gave me the book on it. Now I practice those in a setting just as a doctor or a lawyer would. Then let’s get feedback and talk about this specific practice. I thought that was very professional. (Mark, Group Interview 2, October 30, 2019)

For Mark, the video and the inclusion of specific avenues for professional development made this a more professional process and one that felt more like professional development than evaluation. Mark attributed this to the use of video and the fact that the teacher could go back and review the video as a point of reference. In his interview following the first cycle, he said, “The video provides that, because you can actually now look back at yourself and say, ‘Okay, I
see now. Andrew thinks that I need a little bit of work here. He’s given me multiple methods in order to improve’” (Mark, Interview 2, June 20, 2019). Mark’s comments here focus more specifically on the performative routine, or how the routine was enacted by us, rather than how it was ostensibly designed in its ideal form. While the existing system does not preclude the observer from providing multiple methods of improvement, the Swivl Cloud platform made it easier to link suggestions to specific comments. We will return to this feature in the next theme on linking teacher evaluation to professional development.

During the first version of the routine, we realized that one of the most powerful features of the Swivl Cloud platform is the opportunity for the teacher to dialogue directly with the observer by responding to particular comments and questions. I emphasized and even encouraged this as an option for teachers to take advantage of during the fall cycle. Although the existing teacher evaluation system does afford teachers an opportunity to comment, Jonathan found that not as helpful, because “it’s a comment at the end, and it’s not built into a specific moment in your lesson” (Jonathan, Interview 2, June 19, 2019). Jonathan believed that the ability to respond to individual comments would provide teachers the opportunity to clarify their intention “where there could be misconceptions,” so that such confusion could be “tackled and handled immediately” (Jonathan, Interview 2, June 19, 2019). Based upon this feedback, we made responding to an observer’s comment an optional but highly encouraged part of the routine.

After discussing the positive contribution that this ability to reply to comments would add to the process, the teachers and I collaboratively decided that my comments should be shared with them at least 24 hours ahead of the postobservation. In Jonathan’s words, “This way, I could look at it. Then we can have an honest conversation about the questions that you have.
Then I can provide you with a little bit more feedback on those questions” (Jonathan, Interview 2, June 19, 2019).

Going into the second version of the routine, we looked to increase opportunities to make the experience more conversational and equitable. This included not only sharing the comments at least a day in advance but also encouraging the teachers to reply to individual comments. The participants welcomed this addition. For example, Jonathan thought that it allowed the teacher to be more involved in the process and increased teacher agency: “It’s nice to be able to create a dialogue back and forth using the Swivl. That gives teachers an opportunity to explain a little bit further their intention when they’re teaching these lessons” (Jonathan, Interview 3, October 22, 2019). The emphasis on responding to comments made the second version of the routine seem more professional to me because it gave the teachers an avenue to respond to feedback that I provided them, encouraged further reflection on their own practice, and fostered a more two-way conversation.

At our final group interview, Jonathan was actually surprised at how well it turned out. Without any prompting, he addressed our small group: “I have yet to hear a negative comment from my two colleagues and even coming out of my own mouth about this process. It’s been enjoyable. It’s been calming. It’s made me feel like more of a professional” (Jonathan, Group Interview 2, October 30, 2019). In true professions, the practitioners hold a lot of the knowledge and expertise and are trusted to play an important role in their own evaluation (Mehta, 2013). Similarly, for Jonathan, being treated as a professional meant having a meaningful role in his own evaluation—a role in which his expertise and experience were respected and put to use.

Interestingly, Jonathan came into the process the most skeptical of the three participants, partly because the members of his department were very wary of the study, viewed it as a
“gotcha” tool that administrators would use to remove teachers, and thought that he was foolish to participate. For Jonathan, the permanency of the video and the fact that it could be viewed by more than one administrator was an important factor. Although it was an important factor for Jonathan, he also shared that it was a source of concern for his department when he shared the nature of the study with them. “I was bringing up this study to a couple colleagues. They voiced a lot of concern about having the permanence of a video like this being out there. You don’t want to accidentally say something or accidentally do something” (Jonathan, Interview 2, June 19, 2019). When asked to predict whether he would submit what he considered his best lesson during the fall cycle when it counted as an official observation for his evaluation, Jonathan answered:

Knowing that it’s going to be in towards my evaluation, it’s 100 percent going to be my best. The more you think about the actual study, the more you realize how much more weight a recorded video holds than just an observation in the classroom because it could be evaluated by not just one person but by many. It could be evaluated over a course of time. It could be evaluated with different context, and different people viewing it in a different light or in a different way. (Jonathan, Interview 2, June 19, 2019)

As has been noted, Jonathan became a true believer in the process, especially its ability to make him feel more like a professional than he felt when submitting to the existing teacher evaluation process. For Jonathan, the way the routines ended up being enacted contributed to the professionalism of the process, possibly more so than the way they were originally designed. Although we designed the ostensive routine to promote professionalism through teacher agency, the way that the routine ended up being enacted in practice by Jonathan and me—the performative routine—further amplified the professionalization of the routine, because Jonathan
appreciated the creative ways that he was able to take the routine and came to appreciate my partnership in the process. At our final interview he shared, “[I] loved rubbing it into my colleague today about the process being done having told them about how much better I felt about this observation than I’ve [felt about] . . . the past ones” (Jonathan, Group Interview 2, October 30, 2019). I return to this idea of ostensive versus performative routine in my fourth theme, which focuses on our experience of coconstruction and implementation of the action research. Mark was also overwhelmingly positive about the routine and actually reached out to his department supervisor to request that his second observation of the year to be done through the same process. However, due to the constraints of AchieveNJ, it was required that he have an in-person observation, so we were not able to accommodate his request.

Conclusion

Although the three participants did note some positive professionalizing aspects of the existing teacher evaluation system, they described it mostly negatively, as something done to them, not with them, and an exercise in compliance that did not focus on their professional growth. Overall, all three participants felt an increased sense of teacher agency in the process and, as a result, an increased sense of being treated as a professional during the teacher evaluation process (Biesta & Tedder, 2007). In the next theme, I continue to look at our experiences across both cycles, but this time with a focus on linking teacher evaluation and professional development. Although the increased teacher agency through active participation contributed to an increased sense of being treated as a professional, that same principle contributed positively when I attempted to reframe teacher evaluation as an avenue for professional development. I find that some of the same components that led to an increase in a
sense of being treated as a professional led to improvement in teachers’ practice as they perceived it.

**Teacher Evaluation Has Potential to Support Teacher Learning**

The second goal of this action research study was to explore how the teachers and I experienced the professional development component of teacher evaluation. Over the years, teacher evaluation has had competing purposes—accountability and teacher development—with accountability often being the primary driver (Donaldson and Donaldson, 2012; Marzano et al., 2011; Woulfin et al., 2016). Although our district has tried to focus on professional development and teacher leadership in the district, this same feeling of accountability as primary driver can also be found in Hickory. As Jonathan shared, “I feel over the years, the observation process has become . . . it’s legalized. Everything’s coming from the state, essentially, about how we can and how we can’t observe teachers. It belittles the entire profession” (Jonathan, Group Interview 2, October 30, 2019). The teachers and I set out to explore these competing purposes in the context of the existing evaluation system and the new evaluation routines that we constructed.

At the outset of the study I shared with the teachers that I was interested in designing a routine with them that leaned more heavily toward teacher development than toward accountability. This resonated with the teachers, who felt that such a system would be welcomed by themselves and their colleagues. For Jonathan, it was important to frame the teacher observation process as a method of professional development rather than evaluation:

> You say that you’re going to use this technology to help that teacher in that area. It’s not about necessarily evaluation as it is professional growth. Then you change the mind-set of the entire process. Doesn’t become a tool to make sure that the teacher gets to keep their job. It turns into a tool where a teacher and the administrator are working
collaboratively to grow a teacher in a particular area where the teacher has identified as
an area of need for them. Then it becomes a worthwhile activity. (Jonathan, Interview 2,
June 19, 2019)

He believed this was important because it would help teachers and administrators come to the
observation with a different mind-set, one focused on growth rather than accountability.

One proponent of such a reframing was John Papay (2012), who argued that “if teacher
evaluation is to improve student learning systematically, it must be used as a tool to promote
continued teacher development” (p. 124). He proposes a reframing of teacher evaluation from
measurement to a professional development tool with the goal of improving “instruction by
developing teachers’ instructional capacity and effectiveness. . . . The evaluation system can and
should be seen as a professional development tool and should be evaluated on its ability to raise
instructional proficiency and student learning” (p. 133). As we set out to conduct this action
research, I was interested in how the existing teacher evaluation system promoted professional
development. Subsequently, I designed a new routine based on principles of effective
professional development and then reflected on how the routine promoted teacher growth by
linking teacher evaluation to professional development. Jonathan put it well:

When you institute a program like this, that gives teachers a feeling that their
administrator cares about their growth, and that they want to work with them and not they
just want to observe them to write it down on a piece of paper and say that they’ve
accomplished a meaningful goal for the state. Everybody wins. I think the whole
organization benefits from them. (Jonathan, Group Interview 2, October 30, 2019)

In the next few sections, I identify specific subthemes that more fully examine what I have
learned about the possibilities and limitations of teacher evaluation as a means to support teacher
learning. Specifically, I explore how the existing evaluation system offered limited opportunities for teacher learning, and then I describe our experience with the new routine, making it explicit which feature that promoted professional development could be attributed to either the ostensive or performative routine.

**Existing System: Limited Opportunities for Teacher Learning**

Along with determining how the teachers experienced these routines in light of being treated as professionals, I also set out to document their experience with the routine and its ability to promote teacher learning. Similar to the way the teachers found the existing evaluation system to be deprofessionalizing while containing some positive components, they also found that overall the existing system does not help them to become better teachers—with some exceptions.

**Not conducive to teacher growth: “It doesn’t show them teaching.”**  Although Mark found the standards to be a strength of the existing system, he did not believe that by themselves they were enough to contribute to his growth as a teacher: “The current system provides descriptions of the teacher and their practices; it even allows for the teacher’s selection of data and philosophy, but it doesn’t show them teaching” (Mark, Interview 1, June 6, 2019). Mark also did not see connections being made between his observations and his other avenues of professional development, such as the coaching he was participating in or the professional development plan (PDP) that he developed collaboratively with his supervisor. Although Mark believed that the teacher evaluation system had increased the professionalization of teaching, he did not believe that the existing system was designed to improve teacher practice:

I don’t think the current system would help teach someone else to be a better teacher. It’s informative but not instructive. The current system only gives a location as to where they
are now as a teacher. Obviously, that’s why I signed up for this study. (Mark, Survey, May, 27, 2019)

Overall, Mark did not view the evaluation system as a tool that helped him become a better teacher. Elise also shared that the existing evaluation system affected her practice little. When specifically asked if the current observation system affected teacher development, Elise did not express enthusiasm:

No, I don’t think so. Maybe for the 10 seconds people look at what it says. I know that I spend more than 10 seconds, but I don’t spend a lot of time. I process it and hopefully I implement the suggestions that are made. (Elise, Interview 1, June 5, 2019)

Despite an overall negative opinion on the ability of observations to improve teacher practice, Elise was able to provide a few examples of times when a supervisor provided her with comments that were constructive and led to changes in her own practice.

**Targeted support:** “Actually gave me something to work at.” Jonathan believed that the existing evaluation system could improve teacher practice depending upon its implementation. For example, he believed a strength of the current evaluation system was the opportunities for one-on-one discussions with his supervisor that the pre- and postconferences and end-of-the-school-year conversations afforded him. He found that those discussions allowed him to explain his decisions as a teacher. He used to find these meetings a source of stress as a nontenured teacher but has now grown to embrace them as opportunities for professional growth and development. He believes that these have become a valuable opportunity for growth because they have been among of the few times he has been able to sit down one-on-one with his immediate supervisor and talk about his teaching practice. However, at the outset of our study he thought that overall the focus of the observations was too broad and that it would be
beneficial to narrow the scope to a few indicators. Furthermore, he felt strongly that the supervisors should provide specific feedback and tangible ways for the teacher to improve in areas needing growth. During our first interview, he shared”

I believe there should be examples of how teachers can develop in areas of weakness as provided by the evaluator. For example, if a teacher received low marks in instructional planning, then the evaluator should provide some type of learning experience that could help that teacher develop in the area. (Jonathan, Interview 1, June 6, 2019)

During multiple interviews over the course of the study, he kept coming back to this idea that for teacher practice to improve through the observation and evaluation process the onus should be on the evaluator to guide and support the teacher to help them grow their practice. To illustrate this point he told the story of the most useful observation he had.

The most valuable observation I’ve ever had in 12 years of teaching was when [my supervisor] gave me a book on in-class questioning. Nobody had ever given me a resource. They’ve only given me comments in my entire career. Finally, somebody was like, “Here, Jonathan, I noticed you were struggling a little bit here with things like wait time, and maybe proper phrasing in terms for questioning, to be a little more inviting, etc.” I went through that. This is when I was nontenured. I went through that. I looked at my lesson going forward. I made a couple of adjustments. I sent her some feedback on that as well as the reference she gave me, and it totally changed the outcome of my performance for the remainder of the year and beyond. It was the only time that I’d ever received helpful information in an observation that actually gave me something that I could work at. (Jonathan, Group Interview 1, June 24, 2019)
This quote shows how important it is for Jonathan that the teacher evaluation system have a connection to his development as a professional. He valued this observation because the administrator provided him with a resource that could aid his teacher practice. Based upon this feedback from Jonathan, I made certain to provide him resources that would help his development. This led me to discover a feature of Swivl Cloud that we emphasized as part of the second version of the routine.

The participants had mixed feelings about the existing evaluation system’s ability to promote teacher development. I set out to design an organizational routine explicitly aligned with components of effective professional development (Ball & Cohen, 1999; Darling-Hammond et al., 2017; Hawley & Valli, 1999; Klein & Taylor, 2017). In the next section I describe where and when the participants and I experienced aspects of this ostensive routine that promoted teacher development.

**Ostensive Routine Supports Active Learning**

The organizational routine that I designed for the first cycle included multiple components of effective professional development as defined by Darling-Hammond et al. (2017). First and foremost, it focused heavily on feedback and reflection by providing a vehicle for observers to give time-stamped, specific feedback to teachers linked to actual teaching while also providing an opportunity for teachers to watch themselves teach and reflect upon that teaching. The routine was designed to incorporate active learning due to the teacher’s role in the process, was content focused in that it focused on the actual instructional delivery in the context of the classroom, and provided coaching, with the observer being “rebranded” as a coach providing support.
Effective professional development often includes active rather than passive learning (Darling-Hammond et al., 2017). However, as described earlier, in the existing evaluation system, the participants described a process in which they felt that actions were being done to them rather than allowing them to engage as an active participant. In the new routine, after recording their lesson or lessons, the teachers had the opportunity to choose which one they wanted to submit for their observation. Having this opportunity allowed them to play a more active role in their own evaluation (Klein & Taylor, 2017). For example, once she was able to get past the awkwardness of watching herself on video, Elise found herself engaged in self-evaluation and reflection:

I do remember noticing things that were happening in student engagement that I wasn’t aware of during the lesson. It was catching some things that I wasn’t seeing and that was helpful. That was valuable to me to make some changes. (Elise, Group Interview 1, June 24, 2019).

Because the new routine made her an active participant, she was able to notice practices that she would not have been privy to without the use of video and the expectation that she view the videos critically in order to choose one for submission. During the first cycle, Jonathan recorded eight videos and “watched four or five of them” (Jonathan, Interview 2, June 19, 2019). “It was strange,” he said,

but fun. It’s funny watching yourself on video teaching. It’s something I’ve never done before. It was eye-opening to notice a lot of the gestures that I make with my hands. I didn’t realize how active I was in using my hands when I teach and how animated I feel like I can be. . . . A lot of times I try to elicit some type of reaction by joking with them and having some light humor thrown into the mix when I’m teaching a lesson. I didn’t
realize how much I actually was doing that until I watched it back. I definitely thought a little bit more critically about some of the reasoning behind why I do what I do after having watched the videos. (Jonathan, Interview 2, June 19, 2019)

He explained that he noticed a lot of mannerisms of which he was not aware. Similar to Elise, he initially focused on superficial aspects of the recording such as appearance, voice, and mannerisms. During the first cycle, Jonathan became a little more comfortable watching himself teach on video, which he had never done before. However, he said, “I also noticed instances where I could’ve asked follow-up questions, but instead gave advice that the student could have come to on their own” (Jonathan, Interview 2, June 19, 2019).

Mark took a slightly different approach than both Jonathan and Elise. Although he taped himself 12 times, he ended up only watching the one video that he decided to submit. This was not his original plan, and he shared, “The hardest part was getting myself to [watch] the video” (Mark, Interview 2, June 20, 2019). He observed, “I was surprised that I didn’t want to go back and look at myself on video” (Mark, Group Interview 1, June 24, 2019).

I didn’t expect that I’d watch all of them, but after having filmed for 11 lessons, I felt like I had a pretty good sense of the rhythm of each one. I could have chosen any one. I would have been comfortable with any one. (Mark, Interview 2, June 20, 2019)

Even though he didn’t watch all 11 lessons, Mark noted, “I spent much more time polishing [the lesson] than I would have [if] it was an ordinary administrative observation” (Mark, Interview 2, June 20, 2019). These are all examples of the teachers playing a more active role in own evaluation, studying their own practice and pointing out areas for improvement and thereby developing a stance of inquiry (Ball & Cohen, 1999; Cochran-Smith & Lytle, 1999). All of this occurred before they even submitted the lesson to me to review and provide feedback. This was
one of the most powerful aspects of the routine as it was designed. By giving the teachers the freedom to record multiple lesson and choose one for submission, I increased the chances that they would be actively involved in their own evaluation, because this process promoted reflection and self-study, which is all but absent in the existing evaluation system. I found this process promoted teachers studying their own practice and developing a greater stance of inquiry. This was by design.

The final part of the routine was the postobservation conference. In the traditional postobservation conference, the teacher and supervisor might have a conversation about the lesson in question. This postobservation is supposed to take place within five days of the actual observations. Sometimes the supervisor will share their observation report with the teacher in advance of the postobservation; other times the teacher does not see the report until after the postobservation. For our routine, I shared the report ahead of time, with all of the comments. Then during the postobservation, the teacher and I watched the video together, focusing on the comments I had included ahead of time. For Mark the collaborative viewing of the video was extremely powerful: “It makes you aware of things that you weren’t necessarily aware of. That’s the beauty of the video” (Mark, Interview 3, October 25, 2019). Jonathan also thought that viewing the video collaboratively was a powerful aspect of the routine. He appreciated collaboratively viewing of the lesson because it provided him a chance to explain parts of the lesson when and where appropriate, because not everything can be seen or heard by the Swivl. He shared, “You had asked me a question about giving student feedback. One thing that you didn’t see on the video because the video was fixated on the presentation was the rubric that I had at the desk with me” (Jonathan, Interview 2, June 19, 2019). He connected the use of video to his work as a coach with his players:
Although it’s stated in paper, it’s not as clear to you as when you’ve watched it on video. We’ve had conversations like this. It’s the same when you tell a kid in lacrosse practice, “You’re doing this wrong,” and then you show them the video and they’re like, “Oh, I get it.” It’s the same principles. It’s why you do film in sports. It’s because you want kids to grow and get better. Teachers can do that. (Jonathan, Group Interview 1, June 24, 2019)

Mark also thought the sequence of recording the videos and then completing a reflection before sending one of the videos to the supervisor was effective. He appreciated being able to proactively address anything that he thought did not go as well as he would have liked or a decision that he would have made differently. Furthermore, he believed that this preparation made the process “a pretty good collaboration” and leveled the playing field between teacher and supervisor (Mark, Group Interview 1, June 24, 2019).

In designing the routine, I purposely included a collaborative viewing in which the observer would play the role of coach. The thinking was that this would align with the coaching that has been found to be effective in promoting teacher growth (Darling-Hammond et al., 2017). The participants found the collaborative viewing useful, but they did have suggestions for how it could be improved based on how we performed the routine during the first cycle. I revisit this in a subsequent section on the performative routine.

It is well documented that opportunities for teachers to receive feedback, reflect on their practice, and then put that feedback into practice are a hallmark of effective professional development (Darling-Hammond et al., 2017). All the participants said that the written feedback aligned to professional standards was one of the few positives of the existing system. However, I
still thought that there was room for improvement, because I looked to take advantage of the
time-stamped feature in Swivl Cloud. For the first version of the new routine, the teachers would
still be getting written feedback aligned to standards, but now it would be tied to a moment in the
video. Below is an image of a comment being tied to multiple standards. In the current system,
this is possible, but very cumbersome to do.

Figure 4.1. Example of linking feedback to multiple teacher performance standards.

**Feedback tied to a specific moment: “I don’t have to go dig for it.”** Moreover, the
Swivl Cloud software is designed so that the comments that are made are anchored to a specific
moment and you can skip directly to that moment with one click instead of having to scroll
through the entire video looking for the feedback. Elise thought this feature was very valuable: “I
like this feature that it jumps right to where you are, where you made the comment. I like that
feature. I feel like it streamlines the whole process” (Elise, Interview 2, June 17, 2019). A few
days later during the group interview Elise came back to this point about how beneficial the time
stamp was for making the process more efficient and more effective: “The time stamp is huge. It
goes right to the moment. The comment is right at that moment. In a written observation that’s
lost because half the time you’re not quite sure when that happened” (Elise, Group Interview 1, June 24, 2019).

Teachers also found the ability to rewind, rewatch, and review in light of a comment to be an effective feature of the program. This feature allowed them to examine their own work, a feature of effective professional development (Ball & Cohen, 1999). For example, at one point in Elise’s submitted video, I inserted a question about a topic she was discussing with the students that was confusing to me. “I think it gives the opportunity to clarify, like if you didn’t understand something, then I could back it up and [explain]” (Elise, Interview 2, June 17, 2019). When asked to compare time-stamped comments to the traditional method of receiving feedback after an observation, Elise responded, “The other way seems disconnected, like I have to try to remember. They saw this in the classroom, they made notice about it, but here it’s very connected in living color reality” (Elise, Interview 2, June 17, 2019). This difficulty remembering can be found in other studies comparing traditional evaluation with video-enhanced evaluations (CEPR, 2014). Overall, Elise found this method of receiving feedback to be “more effective because . . . you see exactly what is being spoken about, and you can see it in living color. You can go back, and see it, and watch it again, and listen to it again” (Elise, Interview 2, June 17, 2019). In a later interview after the second cycle, Elise commented again on the efficiency and efficacy of receiving the time-stamped comment embedded directly into the video: “What I like about it is that you’re able to time-stamp exactly where you’re saying what you’re saying. It makes it feel very useful. It’s right there. I don’t have to go dig for it” (Elise, Interview 3, October 25, 2019). Below is a screenshot from Swivl Cloud showing how the comments are time-stamped and tied directly to the video clip it is referencing.
Feedback that is specific: “It’s not a general statement.” Jonathan also appreciated the fact that the comments were time-stamped and directly tied to a moment from the lesson. “The biggest thing is that, again, opportunities for me to review the comments and have them directly connected to a piece of a lesson so that it’s not a general statement, which is what you would find in many observations” (Jonathan, Interview 2, June 19, 2019). According to Mark, having the comments tied to the video was an effective way to receive feedback. To a question about how he experienced receiving feedback in this manner, he responded:

It’s excellent . . . it’s able to go to the minute and second and say, “Look at this little piece right here.” I mean, this day and age, we’re getting used to that in other forms. Now, when you’re able to go to that exact moment, it’s like, “Wow!” It’s mind-blowing in a way. (Mark, Interview 2, June 20, 2019)

Later on, Mark expanded further upon this feature: “[It] allowed me to look at the decisions I made at each point. When I look at this video it allows me to look at the sequence of decisions that I made. We can look at the craft of teaching” (Mark, Group Interview 1, June 24, 2019).
For Mark the fact that the comments were tied to a particular moment allowed for a deeper analysis of teaching and was therefore a more powerful tool for professional development:

It’s a powerful tool. What we were doing is we’re going deeper into what was happening in the classroom. [Andrew] was actually looking at the word choice that I was using to close a lesson. The way I’ve closed the lesson in my verbal pattern was something we talked about which has never happened in my career and it’s interesting. (Mark, Group Interview 1, June 24, 2019)

![Figure 4.3. Example of time-stamped comment from administrator in Swivl Cloud.](June13.period9)

This specificity was important for teachers looking to improve their practice. They did not find the general statements that were common in the existing system conducive to promoting their growth. As Ball and Cohen (1999) argued, “Teaching occurs in particulars—particular students interacting with particular teachers over particular ideas in particular circumstances” (p. 10). The use of time-stamped comments directly tied to a particular moment is perfect for this requirement of effective professional development.
As the observer in this process, I greatly appreciated the ability to provide comments that I considered more meaningful, specific, and useful than those I usually am able to provide within the existing system. In that system, I spend a great deal of mental energy and time contextualizing my comments so that when the teacher reads them they will be able to tie it to a specific moment in the lesson. However, I am not certain that I am always effective in this practice, and difficulty remembering the particulars of a lesson can be a problem for teachers (CEPR, 2014). In the new routine I developed, I did not have to worry about describing what was happening because it was all right there with the video. This allowed me to focus on providing more useful and analytical feedback and allowed the teachers to see the exact moment.

The second aspect of the use of video that I found to be really effective as an observer was the ability to rewind. Although it may seem obvious, it cannot be overstated how helpful this feature of the Swivl Cloud platform was for me as an observer. I found myself using this feature during every one of my observations. It allowed me to watch the same scene several times and offer more thoughtful comments. Furthermore, if I was not sure exactly what was happening or what was said, I was able to go back and review the video for clarification. This missing of action or dialogue happens more often than we care to admit during a live observation, due to a variety of factors, including the immense number of things happening in a classroom at any given moment as well as the distraction of writing down evidence while trying to pay attention to everything that is happening. Conducting an in-class observation in real time is truly a challenge that requires sophisticated multitasking skills to do well.

As referenced in the previous paragraph, conducting an in-class observation in the existing system is a challenging task requiring high levels of observational and recording skills.
In our current observation system, the observer is asked to critically observe a classroom of a teacher and their students, describe in detail what is happening, attach what is happening to an indicator of teacher practice, and provide suggestions for growth, when appropriate. I found that the use of video made this task a little less challenging by lessening the burden of describing what’s happening. By having the video as the anchor to which I could attach my comments, I was relieved of the task of describing what was happening at that time and place. This allowed me to focus my attention and efforts on a deeper analysis of teacher practice.

To aid this deeper analysis, I found it extremely beneficial to have reference documents available to refer to as I completed the observation. Although the indicators from the Stronge model were embedded into the Swivl Cloud platform, as they are embedded into our current platform, I enjoyed being able to print out other documents and look at them while viewing the lesson. In our district we have several key documents that we have developed and that are used to guide teaching and learning. For example, we have developed a Profile of a Hickory Graduate, which outlines the key skills, competencies, and dispositions that we hope all our graduates will have. All these competencies include clear definitions, and we have been able to create rubrics for the competencies referred to as the 4Cs: creativity, collaboration, communication, and critical thinking. When I was preparing to observe Jonathan during the second cycle, we decided to focus on the teaching of critical thinking, and I found it extremely helpful to have these documents printed out and easily accessible during the observation. Although these documents are available electronically, I have found it clumsy and difficult to toggle between tabs while conducting an observation in the existing system. I could print them out and bring them to the classroom during an observation, but that could be awkward and not always practical, because I often have limited space while conducting an in-class observation.
And, of course, I could have these documents printed and available when I later write a report on an in-class observation. However, I have done this and do not believe it is as effective as having them available while actually viewing the lesson.

**The power of a narrow focus:** “It became more about an ability to refine your practice.” Another aspect of the routine that was consciously designed to be aligned with effective professional development was the narrow focus of the observations. Teaching takes place in particulars, and improvement to teacher practice is most often found in specific decisions and situations, not grand theory (Ball & Cohen, 1999). As noted earlier, the teachers did not find the general statements of the existing teacher evaluation system to be particularly helpful in improving their practice. In my experience, these general statements are partly a consequence of the overall breadth of the evaluation system. In a short time, the observer is expected to note and comment on many indicators of teacher performance. In practice, this expectation often results in a superficial observation that mentions many of the indicators with very little substance behind any of the comments. Therefore, in designing the new routine, I set committed to a narrow focus. It is worth noting that this was one aspect of the routine that did not rely on the use of video. This was accomplished by having the teachers think about and propose a few indicators that they wanted to focus on for the observation. This was collaboratively agreed upon during the preobservation conference. At the group interview at the end of the first cycle, we had an interesting exchange about what area they decided to focus on—and the teachers predicted what they would choose and also what most of their colleagues would choose if provided the same scenario. The conversation started with the participants saying that they and most of their colleagues would most likely pick an area that they were very comfortable with, because they wanted to show off their best for their observation. However, as
we talked, the teachers and I came to see that giving teachers choice, framing the observation process as a learning exercise, and providing a long window for completion would increase the probability that a teacher might actually choose an area that they truly believed they could develop. The following conversation illustrates this evolution of thinking:

Jonathan: When you get to set the parameter for what your improvement needs to be, it gives people a little more choice and power and they might actually give in there.

Andrew: I think that’s what I’m hoping, because you have the agency and you have all the control, that you’ll pick an area that is less than your number one confidence area. You might not pick something that you know you’re going to kill but you have six weeks to work on conferencing and you’re going to get a really . . .

Elise: That’s what I was thinking, by then if you know that’s your focus, if that was your weakest, then for six weeks you’re going to work at conferencing and so by the end of it . . .

Mark: It’s true.

Jonathan: I think a tenured teacher . . . who had six weeks to work on conferencing, even if you send Andrew in a lesson on conferencing and it still wasn’t that great, I don’t think any teacher is going to sit here and think, “They are going to bring me up on tenure charges [laughs] because I didn’t do as well as I should have when it came to conferencing even though I worked on it for six weeks.” I think in reality it becomes more conversational and more collaborative between you and the administrator. (Group Interview 1, June 24, 2019)

This became an aspect of the routine that the teachers found to benefit teacher learning.
At this same group interview following the first cycle, Mark shared, “One thing that I really like about this project … it’s teacher’s choice, it’s teacher-driven, it’s small, and it’s something like, it’s very tangible, it’s very easy for us to look at” (Mark, Group Interview 1, June 24, 2019). Later, after the second cycle was complete, Jonathan echoed the importance of the narrow focus: “The idea of having a particular focus, making it more of a conversation than an actual observation, it became more about an ability to refine your practice. That’s what I like about the focus” (Jonathan, Group Interview 2, October 30, 2019). Elise found the process more effective in targeting specific teaching decisions that could have been made. For example, during the first cycle I commented that it might have been helpful to have an actual map when referring to a region of the world, because the students showed difficulty with this. In our postobservation Elise shared:

[I] thought about the fact that a map would have been a much better way to help them visualize what we were looking at. That would be the decision now, to hang a map, so that we have that going forward. Yeah, that’s going to happen. (Elise, Interview 2, June 17, 2019)

It is important to note that this focus on a narrow set of teacher practices is not something that is necessarily unique to this routine we have developed. In the existing system, too, the teacher and administrator can decide upon a narrow focus for an observation. However, this is rarely done because the observation is usually seen as one of two main opportunities of evidence collection for the final summative evaluation. In other words, due to the way that the existing system is set up, it behooves a teacher to have as many possible pieces of evidence listed in their observation so that at the time of their summative rating there is a large body of evidence to justify a high score. This leads to quantity over quality. For some reason, the teachers in this
Possibly because of the narrower focus, the teachers reported that they carried forward the comments and conversations into their practice more than when they were participating in the observation cycle in the existing system. Mark explained how using video to work on a particular area helped him improve: “That’s the improvement that I’m looking for. That’s the improvement that the video is going to give me. That’s something I’ve been working on in the months since” (Mark, Interview 3, October 25, 2019). Mark expanded upon this during the group interview a few days later: “I feel like the process helped me grow. It’s like I was able to look back with the time stamps to what I was doing. . . . I’ve tried this, and it’s worked a little better. That’s why I like this whole thing” (Mark, Group Interview 2, October 30, 2019). The teachers found the feedback was much more targeted and digestible in this format. “I can’t always remember everything that happens. But the time stamp is right there. That’s really helpful. It’s almost like immediate feedback. We know that you’re talking about” (Elise, Group Interview 2, October 30, 2019).

As an observer, I greatly appreciated the narrow focus. One of the most challenging aspects of the existing evaluation system is the sheer scope of the standards and indicators. In the Stronge Evaluation System, which our district uses, there are six standards and more than 60 indicators. It is impossible to make specific, valuable, and meaningful comments about that many indicators of practice. Although that is not the expectation, in reality I try to hit dozens of the indicators during a lesson. As a result, I am typing furiously at my computer, jumping between screens on my laptop and often missing what is actually happening in the classroom because I am consumed with documentation. The narrow focus allowed me to view the lesson
through a specific lens. Therefore, I believe I was able to provide better feedback to the teacher. Knowing the focus of the observation before it began allowed me to have key documents available if I needed to reference them, and it allowed me to pause the video allowed me to document my thinking without missing what was happening in the classroom. I walked away from these lessons feeling like I had a better handle on what actually happened in the classroom.

**Depth over breadth: “I don’t think you need to watch as many.”** Going into the study, we were aware that you could insert time-stamped comments and rewind the video to rewatch when needed. During the first version of the routine, I made about 20 comments on each teacher’s lesson. During our collaborative viewing, we discussed each of these comments, stopping at each point in the lesson and watching the action that the comment was attached to. During the follow-up interview after the postobservation, Jonathan suggested that, going forward, we focus on fewer video clips during the postobservation conference, which would allow for a deeper analysis of and conversation about those particular moments. He argued that through the comments that they received beforehand, the “teacher understands what you’re looking for… I don’t think you need to watch as many” (Jonathan, Interview 2, June 19, 2019). He thought the number of comments was appropriate and helpful but that it was not necessary to discuss all of them in the same detail during the postconference. For the second version of the routine, he suggested focusing on two to three points in the lesson to watch collaboratively rather than reviewing all the comments made throughout the lesson. The teachers believed that focusing on fewer scenes was a positive change for the postobservation. They shared that this focus allowed the postobservation to attend to the most salient points that could be addressed with the goal of promoting teaching learning. Elise appreciated the deeper dive that focusing on a few scenes allowed:
You want to talk about the areas that you should grow things, those things that you identify this, like wait time, things like that. I want to talk about that, because those are the areas that I want to improve in. It seems reasonable that we would spend some time talking about that. Being very direct with not looking at the whole thing. Do we need to look at the whole thing? I don’t think so. Nice job. Now, these two things. That’s what I expect. (Elise, Group Interview 2, October 30, 2019)

The teachers also clearly believed that the administrator should be mindful of not only pointing out areas for improvement but also areas of strength that the observer noticed. Jonathan referred to this as a “glow and a grow.” I agreed with the teachers and also recognized this as a great learning moment for me as a user of this technology for supervision. While I was not able to name the problem, like Jonathan, I felt rushed while reviewing all the short video clips during the postobservation conference. Therefore, this change to the routine was included in the second version enacted during the fall. Although none of the teachers had much to say about this change after going through the second version of the routine, this was one of the most significant changes for me—the observer. This last aspect of the routine was only part of the second version, because it was a revision suggested by Jonathan after he reflected upon the first cycle. He suggested focusing on a few scenes for collaborative viewing during the postobservation rather than reviewing all the comments and associated clips. This turned out to be a very positive change for me as the observer. In the Group Interview 2, I shared with the group that “I liked focusing on two scenes, watching them with you, and then going through. What I need to be mindful of, and a trap an observer can fall into, is honing in only on the things that need to be fixed” (Andrew, Group Interview 1, June 24, 2019). The teachers agreed with that, noting that it
would be important for the observer to not only point out areas for growth but also pay attention to and highlight good teaching practices.

**Performative Routine Showed Power of Flexibility and Unknown Features of Technology**

Although the ostensive routine was designed to promote teacher growth by following the principles of effective professional development, some powerful components and approaches to completing the routine arose as we performed it. In the following sections I present where and when effective professional development can be attributed to how we performed the routine, rather than solely how it was designed. This includes not only how the teachers chose to enact it but also how I chose to enact it.

**Power of flexibility for teacher.** During the second cycle, the teachers had a longer window to film themselves and choose a video for submission. Based on the teachers’ feedback, we extended the window from two weeks in the spring cycle to five weeks in the fall cycle. No approach was mandated, because we wanted the ostensive routine of the second cycle to be flexible for the teacher. However, I envisioned the teachers following a cycle of recording a lesson followed by viewing the lesson while reflecting critically on and appraising their practice. Based on their reflection and appraisal, they would then look to film another lesson based upon the takeaways from the first viewing. I imagined this cycle would continue until the teacher was satisfied enough to submit their video. What actually happened when the teachers performed the routine was that each of the participants took a different approach based on the flexibility that this longer window provided them. In this way, the flexibility built into the ostensive routine fostered innovative and creative approaches in practice.
It was interesting to see how all three teachers performed the routine differently, and in a way that made the most sense for them. Elise approached it the way that I anticipated all of the teachers would. She took advantage of this longer window to film more than 10 lessons.

I filmed more. We had more time. I filmed nine that I saved [and] a couple I deleted. I filmed in a couple of different settings, because I think I just wanted to see which one I thought was the best. The time allowed me to film more. . . . I liked that I had time to redo the same thing with different groups and perfect it each time.” (Elise, Group Interview 2, October 30, 2019)

Although the lessons varied in length, Elise watched all of them from beginning to end. She reported that she knew that some of them were “duds” right away, but she still watched the entire video even though it was painful.

Mark’s experience with the second version of the routine differed from his experience with the first version mainly due to the extra time provided. Although he intended to film and view multiple lessons during the six weeks, he only filmed one lesson during the second cycle, as opposed to the 12 lessons he filmed during the first cycle in the spring. “I filmed 12 in the spring. I filmed one this time. I didn’t feel like I needed it. That’s part of the beauty of it also” (Mark, Group Interview 2, October 30, 2019). In the spring, Mark filmed 12 lessons during two weeks. At the end of those two weeks he found that he didn’t want to watch himself on video and ended up only watching the one lesson that he submitted for his observation. For this second cycle, he taped himself once and didn’t watch it until three weeks later. In explaining his process, Mark shared,

I didn’t really want to go back and watch the 12 in the spring. Going back and watching this one was different because I waited a month. I went back and was like it wasn’t as
good as I felt about it at the time. . . . There were some rough spots, but that’s what
happens in an ordinary class. I thought it was true. I thought it was authentic. (Mark,
Interview 3, October 25, 2019)

Although he only filmed one lesson during the longer window, the longer window did
allow him to have a month’s buffer in between the time of filming and the time of his
viewing. Mark explained how he came to see this:

I think it was a positive, because you come back that month later and you look at it with
fresh eyes. You have a remembrance of it, but studies always say how you can
misremember things and they don’t go the way you planned. It wasn’t the way you
remember it. Going back a month later and looking at it like, “It wasn’t really as good as
I felt it was.” (Mark, Interview 3, October 25, 2019)

In trying to determine why he ended up filming only one lesson during the six-week
window, Mark thought it might have had something to do with the fact that the window occurred
during the first six weeks of the school year.

It’s tough because this was the beginning of the school year. It’s hard to get into the flow
of things, where you are in the curriculum, and get things set and routines, stuff like that.
Maybe that played a part in not filming 12 of them. (Mark, Interview 3, October 25,
2019)

He came to realize that the longer window allowed him to be observed in September for
the first time in his career. Jonathan also liked the fact that the window occurred at the
beginning of the year, because he found it helped him tie the focus of his observation to district
and building goals that are frequently emphasized during these first six weeks of the school year.
Jonathan took a different approach to filming than either Elise or Mark: “I figured out what lesson I wanted to be filmed and waited for my courses to get to those lessons naturally” (Jonathan, Interview 3, October 22, 2019). He found this to be a real positive, because “sometimes when you have an observation you force a lesson too early or you’re waiting for it because you know an administrator is coming in. . . . Knowing that I had the time frame to work with, I naturally got my students prepared for the lessons that we were going to be recording” (Jonathan, Group Interview 2, October 30, 2019). Whereas Elise watched each lesson after she filmed it, Jonathan filmed all six lessons before sitting down to watch them all, which he did in one sitting. And whereas Mark did not watch his lesson until a month after filming it, Jonathan filmed his lessons over the course of two days and then watched them the following day.

All three teachers found the extended window of the second version to be preferable to the shorter window of the first version, because the increased time allowed each of them to perform the routine in their own way. Jonathan put it well: “It was great. I had the time frame. It was a relaxing process. It wasn’t stressful. That’s why I like it. When it comes down to the time frame and also the filming itself, it’s very teacher friendly. There’s no real stressful component to it. You can control the narrative and the time line in your own way. It made it easy” (Jonathan, Group Interview 2, October 30, 2019). Although the new routine was designed to be flexible and with principles of effective professional development in mind, the real power of this part of the routine was how the individual participants decide to perform it.

**Power of flexibility for observer.** Practically, one of the best parts of this routine was the flexibility that it afforded me as the observer. Because I was not tied to a teacher’s schedule and could access the lesson via my laptop, I could conduct this observation at various times and in various locations. For example, during this action research study I conducted some of these
observations at my desk in my office, but also was able to watch the videos and provide feedback while my one-year-old daughter napped and in the early morning hours at my kitchen table while my daughter was still sleeping. In my current role in the central office, it was very convenient to be able to observe teachers in three different schools without leaving my office. Furthermore, I was able to begin an observation, pause it, and then continue it later. This flexibility was extremely helpful for me as the observer. On a practical level it allowed me to begin an observation and then pause if a pressing concern arose that needed my immediate attention. It also provided substantive benefits, because I found myself pausing a video when I didn’t feel that I was in the right frame of mind for conducting a fair and thorough observation. I paused the video, walked away, and came back to it the next morning. Without video, that would not have been possible. Of course, this is not to say that in-person classroom observations are not important. However, the flexibility was important to my work as a leader in being able to provide high-quality feedback. In my previous role as a building principal, I would have appreciated the flexibility even more, because the demands on my time were even greater in that position.

The increased time spent on observations is a common complaint of administrators in the wake of the extensive requirements of AchieveNJ. It is important to note that I did not find that this routine reduced the amount of time that I spent on teacher evaluation. I could see myself becoming more efficient at the process, because this was also a new routine for me, and that “newness” inevitably slowed down the performing of said routine. However, an administrator could use the routine my fellow study participants and I developed as a time-saving routine if they chose to give the videos a superficial skimming and insert a few comments. As I will discuss further in Chapter 5, this routine is not “administrator-proof,” and there is no guarantee
that an administrator will not use this routine simply as a time-saving tool. However, with a deep understanding of the purpose of the components of the routine, I believe that an administrator would be less likely to use it as described above.

**Technology that led to new opportunities.** As we enacted the routine, we discovered several features of the Swivl Cloud software that became an important part of the second version of the routine. In this section I explore those features and describe how the teachers and I came to experience them when they were integrated into the second version.

During the spring cycle we realized that one of the interesting features of the Swivl Cloud platform is the opportunity for the teacher to respond directly to individual comments and questions. I emphasized and even strongly encouraged such responses as an option for teachers to take advantage of during the fall cycle. Although the teachers had the opportunity to respond to comments during the first version of the cycle, we did not discuss or emphasize this feature of the platform. As we collaborated on the process during the first version, it became apparent both to the teachers and to me that this was a powerful tool that should be utilized better going forward into the next cycle of our action research.

For Elise, the encouragement to reply to individual comments was a welcome addition: “[It] does feel like it gives me a chance to be more part of . . . less under the microscope and more part of what’s happening” (Elise, Group Interview 2, October 30, 2019). Jonathan concurred with Elise’s thought that it allowed the teacher to be more involved in the process and therefore experience an increased in teacher agency: “It’s nice to be able to create a dialogue back and forth using the Swivl. That gives teachers an opportunity to explain a little bit further their intention when they’re teaching these lessons” (Jonathan, Interview 3, October 22, 2019). As we enacted the routine, this change further made them active participants in the process,
evened the playing field between teacher and observer, and promoted the idea of teacher 
evaluation as an opportunity for professional development.

Another aspect of Swivl Cloud that was not known going into the study but ended up proving useful was the ability to link documents and websites directly to comments. Jonathan appreciated this aspect, because he put a lot of importance on the supervisor providing him support in areas where the supervisor made constructive. He particularly liked how the supports can be directly tied to a particular comment rather than to the observation-at-large. The image below is an example of a comment I made that included a link to a blog about the suggestion I was making to Jonathan about his teaching practice.

![Figure 4.4. Example of comment with linked resource.](image)

This feature of Swivl made the feedback more effective because it provided a suggestion that could be followed up on, and responses from the teachers involved in the study, described earlier, suggested that this opportunity to follow up increased the probability that a given comment would help teachers improve their practice by evolving the observation from a one-off experience into an interconnected component of a more sustained professional development approach (Darling-Hammond et al., 2017).
As noted earlier, Elise believed that the existing teacher evaluation system was effective if the purpose was to hold teachers accountable but lacking if the purpose was to help teachers develop and improve their instruction. When posed a similar question, but this time about the effectiveness of the collaboratively developed routine to improve practice, Elise said:

If the purpose is to improve teacher practice, then it’s an A plus, because that’s what happened to me. I started off with my first conversations on a sticky note, my points that we wanted to have. Then I realized after my first sticky note, in the first filming, I was like, “Wait, back it up. By the time I got done I improved my prompting. . . . If improving practice is the goal of observation, then I think this does it. (Elise, Group Interview 2, October 30, 2019)

Overall, the teachers found that this routine that we designed and performed together was helpful in improving their teacher practice. As shown in an earlier section, the teachers also said that the routine increased their sense of being treated as a professional. In the next section, I explored (a) the teachers’ overall experiences of being part of the routine and study, focusing on how, if at all, those experiences changed how they view teacher evaluation and (b) how they experienced partnering with me as a co-researcher in this action research study.

Trust Is Central When New Evaluation Routines Are Being Developed

The fourth and final theme is about how the relational aspects of working collaboratively on the coconstruction of the routine shaped the overall experience for both the teachers and me. In this final theme I describe how participating in an action research process might have increased the level of trust and therefore affected how that process played out. I examine how the teachers came to experience that level of trust and whether specific components of the routine or steps that I took helped build trust. While exploring this idea of relationships as
important in action research, I also explore how I came to further, and more deeply, realize that relationships are important when one conducts teacher evaluations and introduces a new organizational routine.

**Importance of Clear Communication and Transparency: “A Certain Level of Trust Is Created”**

Going into the study, the teachers knew that they would be co-researchers in a doctoral dissertation and that as co-researchers they would have opportunity to reflect on their experiences and coconstruct the routine that would be used in the second cycle. Furthermore, they knew it was possible that the routine we developed would be scaled up across the district in coming years.

Based on their comments, I believe that giving the teachers positions as co-researchers with me helped me win their trust in both me and the process. I attempted to be very transparent about the purpose and process of the study. I believe because I did so, the interactions I had with the teachers were a little different than they might have been if this work had not been conducted as part of a larger research project. In other words, if we were exploring this as an alternative model to supervision but I was not frequently asking their opinion and conducting formal interviews with them, it is possible that they would not have had the same level of trust. Mark alluded to this in the final sentences of the following quote when he referred to the interviews and emails that had been part of the process. At the final group interview, the group also discussed, mostly unprovoked, how participating as a group in the study might have contributed in some way to their high level of trust in the process. The following exchange illustrates the trust that they felt:
Mark: Here’s something I hadn’t thought of. Was it the fact that we’re in a group going through this, did that affect it and that, somehow lent itself to the trust factor?

Jonathan: We knew that there were two other people going through the same process, that we were somewhere in Hickory going through the process that we were going through and having these same conversations with Andrew. (Group Interview 2, October 30, 2019)

Upon reflection, I think some of those procedural aspects of the research study helped foster trust between me and the teachers involved.

During the final group conversation, I asked whether the teachers found it off-putting when I inserted a comment into their video that was outside the scope of the narrow focus that we had collaboratively decided upon. Even though I had told them up front that this was a possibility, I was still curious to hear their reactions. However, the teachers’ responses surprised me as they began organically discussing the trust and relational aspects of the process and how those factors may have had an effect on the process itself. Here is the exchange:

Andrew: Was it off-putting at all that it wasn’t part of our focus?

Mark: No . . . it wasn’t off putting for me to discuss that with you individually, if it was someone else, maybe. I don’t know why that is. Maybe it’s the personal or the professional trust that we have for one another. I know that you’re not out to get me. If someone else, I’m not quite sure. I’m not quite sure about that. Maybe it would be. Maybe I would take offense, but I don’t. I know that within this framework, we are looking to research things and maybe get a little better at something. I don’t know.

Andrew: That’s really interesting.
Mark: Right. Maybe it’s the interviews that we’ve done. You sent the emails out and things like that. A certain level of trust is created. (Group Interview 2, October 30, 2019)

Several interesting things can be teased out of this back-and-forth dialogue with Mark. First, he mentions that he is not threatened by any of my constructive criticism because of our personal and professional trust and that he knows that I am not “out to get” him. He expands on this by saying that part of the reason he infers that I am not out to get him is that we are doing all of this work within the context of the research project.

Teacher as Equal Participant in Evaluation: “You and I Are on the Same Level Playing Field”

The routine that I designed for the first cycle aimed to foster increased teacher agency and enhance the teachers’ perception of being treated as a professional. As we have seen, several aspects of the routine were included explicitly to empower the teacher as an active participant in their own evaluation. The exchange below illustrates how Jonathan experienced these aspects of the first routine. Mark piggybacked off Jonathan’s comment to explain how he found these aspects to be trust building as well.

Jonathan: I like the professionalism of it. That’s where teachers can take some ownership over their own development instead of having it come from the top down. That’s the key factor to any observation for a teacher at our level. It should be based on what the teacher believes that they need the most amount of work on and how does the administrator help them get there. That’s the point of the observations to me, for a tenured teacher.
Mark: To me that establishes a certain level of trust, too. As we had talked about, I know you’re not using the video to come and get me. You and I are on the same level playing field, and we’re working on helping students and enriching the craft. Not only just for ourselves, for others as well. (Group Interview 1, June 24, 2019)

In the literature review, I shared how changes in organizational routines can often face resistance and that those looking to make changes would be wise to include the teachers in designing or revising a new routine (Conley and Enomoto, 2009; Sherer and Spillane, 2011). Knowing this, I designed the process of the study to be transparent and collaborative. From the outset, I shared with the teachers the goal of the study and continually emphasized that they would be coconstructing the second version of the routine and that they would have significant input into the routine that was ultimately presented to the district in the coming years. Mark’s quote above shows how the fact that we were working on something together for the betterment of other teachers and teaching in general helped foster trust and contributed to a strong relationship between him and me.

On several occasions during the study, I clearly took the advice of the teachers as to how to adjust and revise the routine moving forward. School leaders—including me—often feel challenged when asked to take advice from a teacher, because a leader typically has a vision of where they would like the process to end. The following example shows a time during this process when I listened with an open mind, considered the alternative, and decided to follow a teacher’s recommendation that ultimately came to strengthen the routine. Jonathan used the flexibility of the routine to choose bits and pieces of a longer lesson for his final submission. Knowing that the routine limited him to a video submission of about 20 minutes, he took a 55-minute lesson that he taught, and he edited it to include the 20 minutes that he thought best
exemplified his teaching and our agreed-upon focus of teaching critical thinking skills. At first I was surprised and a little alarmed that Jonathan took this approach. I was not sure whether I thought this form of video presentation should be allowed if the video was to be part of the official evaluation process. However, after listening to his rationale and the process he undertook to review his video and splice together the parts he wanted me to see, I believe that he learned as much from the process of cutting, choosing, and putting together the final 20 minutes as he did from any other part of the routine. If the purpose of the new routine was for teachers to analyze their teaching for the purposes of improvement and feel like a professional while doing so, then I would argue that Jonathan’s approach achieves both of those purposes for him. It was actually Jonathan’s approach to this that convinced me that while the routine needs some structures and guidelines, the teachers involved in the routine would be best served by having the maximum flexibility afforded to them.

**Trusting the Observer: “Would I React Differently If Another Administrator Pointed These Things Out? Would I Be Defensive?”**

Through episodes like this, the teachers came to trust that I had good intentions for this work and that I valued their feedback. Due to this, these teachers were open to receiving feedback from me. One such exchange demonstrates this dynamic:

Mark: Can I ask a question on this one? You [Andrew] and I discussed a little bit, but as I’m watching this, I looked at the very specific questions you were asking, particularly about the transitions. My reaction was, “Wow, I don’t mind Andrew pointing these things out to me,” but it’s sort of like you feel how vulnerable you are at a specific moment. It’s like, “Would I react differently if another administrator pointed these things out? Would I be defensive?” I wasn’t defensive with this. Did you guys have that?
Jonathan: That’s an interesting piece of this. I do find Andrew to be a very fair observer comparatively to others. That’s a really interesting point. I don’t know.

Mark: I realize if I had been observed by someone else in the past or whatever, I might react a little differently. I may not take that commentary to heart as I should.

Jonathan: Maybe you’d be mad about it. (Group Interview 2, October 30, 2019)

Several times over the course of multiple interviews during the first version of the routine, the teachers talked about how they believed that both the use of video itself and the collaborative process that we followed leveled the playing field between teacher and administrator by giving the teachers a good deal of power in the process and the routine. In the group interview following the first cycle, I followed up on this idea of leveling the playing field:

Andrew: That idea of leveling the playing field came through a couple times. That was interesting, that it was like the video allows you to hold the administrator accountable, in a way, for his or her comments. It also levels the playing field in that it doesn’t seem punitive or top-down. It was very much like we are looking at this together.

Mark: Exactly.

Andrew: I want to conserve that as much as we can, because I think that is right. If it’s going to be powerful, you have to have that trust. I think maybe some of your colleagues don’t have that trust, at this point, they’re not in this. Mark also said he thinks everyone in the English Department at the middle school should tape something and share it at a meeting. He said, “I’d be the most hated person in my department.”

Jonathan: Yes, I would be too, yeah.
Andrew: It’s interesting, you guys are all into it, but you also have this skepticism of it getting bigger, like the conspiracy theorists coming out, or whatever. (Group Interview 1, June 24, 2019)

Partly due to their participation in the process, these teachers were comfortable with the new routine. However, they knew they might be the exceptions rather than the rule in their respective departments and schools. Though enthusiastic about the routine’s promise, they shared that their colleagues did not have the same positive outlook. I attribute this view of my coparticipants mostly to the fact that these other teachers were not part of the study and therefore did not benefit from the trust created through our collaborative work. In the final chapter, I discuss the role of trust further as I think about what recommendations I would make for my district if we were going to scale it up in future years.

In this chapter, I shared my findings of this action research study that I participated in with a group of three teachers during two cycles. I was interested in documenting the experiences of three teachers who participated in a new organizational routine around teacher observation that I developed. Specifically, I was curious how they experienced the different components of the routine and whether they came to view observation and evaluation differently through their participation in that routine. Furthermore, I was interested in exploring how their experiences contributed to the codevelopment of the second version of the original routine and how the teachers experienced being actively involved in that process. I found and have shared how the three teachers and I had, overall, a very positive experience participating in the original and revised versions of the routine. In the next and final chapter, I discuss these findings in light of existing research and my conceptual framing as well as discuss implications for practice.
CHAPTER FIVE: CONCLUSIONS AND IMPLICATIONS

The research question that drove this study was, “How do we experience the organizational routine that I developed using video for teacher evaluation for improving teacher practice as compared to the existing teacher evaluation system, and how does our experience inform the routine’s continued development?” Over the course of two cycles, I collaborated with three teachers in my district on developing a new routine for teacher observation as part of the larger teacher evaluation system. During the first cycle, we enacted a routine that I developed. At the end of this cycle, we collaborated to revise this routine so that we could use a better version of it to conduct one of the teachers’ two official observations during the 2019–2020 school year. Throughout the process I documented our experiences using both individual and group interviews.

In Chapter 4, I put forth the major findings from my action research study on the coconstruction of a new organizational routine for teacher evaluation. Overall, I found that our new routine increased the teachers’ perceptions of being treated as a professional and also helped them improve their practice as they perceived it. I also found that the relational trust of the coconstruction and collaborative research played a key role in the success of the design and implementation of this new routine. I presented my findings using four themes. The first focused on the existing system: teacher evaluation system as transactional and deprofessionalizing. In this theme I shared how the teachers found the existing evaluation system to be very bureaucratic and not effective in contributing to their professional growth. In the second theme, new routine: cultivating professionalism through teacher agency, I explored how both the first and second versions of the routine increased the participants’ perceptions of being treated as professionals due to increased teacher agency in the teacher evaluation process.
The third theme was teacher evaluation as supporting professional development, where I shared how the teachers and I experienced the routines in terms of improving teacher practice by aligning with the tenets of effective professional development. When possible, I differentiated between effective components that could be attributed to the ostensive routine and those that could be attributed to the performative routine. In the fourth and final theme, trust is central when new evaluation routines are being developed, I examined the relational aspects of the study, specifically how trust affected our experience, how that trust was developed, and what lessons could be learned from the process.

This final chapter includes a discussion of my findings in light of the literature and my conceptual framework, followed by recommendations and implications. The discussion is organized into two main sections. In the first, I revisit the research literature on teaching as a profession, teacher evaluation, and effective professional development. After each of these summaries, I summarize what I have learned and then briefly discuss how what I have learned either connects to or enhances the existing research. The second section focuses on organizational routines. I revisit the research on organizational routines and share my findings, paying special attention to which findings are from the ostensive routine and which can be attributed to the performative aspect of the routine. As a reminder, the ostensive routine is the routine in its idealized form, whereas the performative routine is the routine as it is actually practiced. I conclude this section by considering how my findings connect to the theoretical literature on organizational routines. My first contribution to the literature is that the relational aspect of codesigning a change to an existing organizational routine is essential and can help promote both higher perceptions of professionalism and increased teacher practice. My second contribution to the literature is that action research and organizational routines are a good match
and that the iterative process of action research cycles works well with the concepts of ostensive routines and performative routines. After discussing my findings in light of the existing literature and my theoretical framework, I make recommendations and draw out implications for teacher evaluation, addressing some of these recommendations to school districts and some to state policy makers. I conclude this chapter with a brief final reflection.

**Discussion of Findings**

In this section, I revisit the literature concepts from my conceptual framework, including teaching as a profession, teacher evaluation, effective professional development, and organizational routines. After each of these short summaries of the literature, I briefly report my findings and discuss how they connect to the existing literature.

**Teaching as a Profession**

In my study I set out to see how teachers perceived the teacher evaluation system treating them as a professional. This included their perceptions of the existing teacher evaluation system as well as how their perceptions changed through participation in the first and second versions of the routine that was the heart of this study. Because I was exploring their perceptions of being treated as a professional, it was important to review the literature on teaching as a profession.

From its inception, teaching has had a difficult time establishing itself as a true profession in the United States (Goldstein, 2014; Lortie, 1975; Mehta, 2013). Teaching has been plagued by isolationism and a privacy of practice that have made it difficult for teaching to emerge as a profession with shared standards and a common knowledge base (Mehta, 2013). Furthermore, a profession must have a way to translate knowledge into practice that involves the practitioners playing an active role in the process. While teacher-as-researcher has been proposed as a model (Dewey, 1896), this view has largely lost out to the more bureaucratic view that knowledge of
teaching is something created outside of the classroom by researchers, who then provide this knowledge to teachers as something to implement in this classroom (Mehta, 2013). This disconnect is one of the main factors preventing teaching from breaking through as a true profession. As a result, the conceptualization of teacher learning that has dominated throughout the years has been the knowledge-for-practice view: that knowledge is created outside the classroom and then needs to be implemented in the classroom by teachers, who are not actively part of the knowledge creation. This has come at the expense of the knowledge-in-practice and knowledge-of-practice views, which are more aligned with a professional view of teaching (Cochran-Smith & Lytle, 1999). One of the purposes of the routine developed for this study was to increase teachers’ perception of being treated as a professional as part of teacher evaluation. I also wanted the design of the routine to provide flexibility in the form of choice. The teachers were able to choose how to implement this routine. The participants in my study perceived themselves as being treated as professionals more through participating in this routine than through participating in the existing teacher evaluation system, and they attributed much of this perception to the fact that they played an active role in the process and were given substantial control and choice over how the process unfolded.

**Teacher agency.** All three participants mentioned that being able to choose their own video for submission was a key factor in their increased sense of being treated as a professional. This choice gave them increased control over what their observer would see as part of their official evaluation. Jonathan remarked how having this power gave them control over the process, which made them feel better about the whole process. Mark concurred and added that this control helped them trust the system. Overall, the participants found that the routine changed from something that was being done to them to a process in which they played an
important part. They became researchers of their own practice and, by doing so, were able to use
the teacher evaluation process as a form of practice-based professional development. This
finding agrees with the literature that has found that practice-based professional development
contributes to a higher sense of professionalism (Ball & Cohen, 1998; Cochran-Smith & Lytle,
1999; Lampert, 2009) and the literature that has reported that an increase in teacher agency
during the evaluation process contributes to feelings of being treated as a professional (Biesta &
Tedder, 2007). Through active participation, the teachers can a great deal of control over their
own evaluation, a hallmark of a true profession (Mehta, 2013).

**Self-critique and reflection.** Two of the participants also noted that being able to
choose from a number of lessons and complete a reflection before submitting the video increased
their sense of being perceived as a professional. They found that these aspects of the routine
caused them to reflect upon and critique their own teaching in ways that were not a part of the
existing teacher evaluation system. All teachers found that the process encouraged them to be
more reflective, which aligns with the knowledge-of-practice approach to teacher learning,
which calls upon teachers to develop a stance of inquiry and be researchers of their own practice
(Ball & Cohen, 1999; Cochran-Smith & Lyle, 1999). The teachers in this study seemed to
become researchers of their own practice and shared that they learned the most during this self-
critique and reflection process, before they even hit the submit button. They shared that they felt
that they became their own teachers throughout the process by watching multiple lessons,
critiquing themselves along the way, and improving their teaching practice more than they ever
would have during the existing routine. As stated above, they became researchers of their own
practice and developed stronger reflection and analytical skills in that role as researcher (Klein &
Taylor, 2017; Rhine & Bryant, 2007; Sherin and van Es, 2009).
Collaborative viewing. Another aspect of the routine that was designed to increase teachers’ agency, and therefore their sense of being treated as a professional, was the collaborative viewing of scenes during the postobservation (Cochran-Smith & Lytle, 1999). However, as we enacted the first version of the routine, this collaboration did not have the intended results, partially due to the number of scenes we viewed together and my approach to viewing these. During the first cycle, I made roughly 20 comments on each teacher’s video submission. At the postobservation conference, we discussed all 20 comments using the playback feature on the online platform. This resulted in a more superficial and rushed review of the comments. This tendency to favor quantity over quality appears in many aspects of the current evaluation system. As presented in Chapter 4, the teachers and I found that narrowing the focus of these observations for the second cycle helped make the routine much more effective in improving teacher practice. This contrasts with the existing evaluation system, in which a teacher is incentivized to provide data on as many indicators as possible so that at the time of their summative evaluation they have a strong case for being rated highly. In the upcoming section on recommendations for teacher evaluation, I propose a modification to the existing system to help resolve this tension.

As we began to revise the routine, Jonathan suggested that in the postobservation conference we discuss fewer scenes, in more depth. As a result, during the second cycle, I made about the same number of comments but chose only two or three to watch collaboratively during the postobservation. This allowed for a less rushed viewing and thus more conversation between me and the teacher. I considered this revision to be one of the best improvements from the first to second version of the routine. I listened to one of the teachers, was persuaded by their
argument, and therefore changed the routine. This change not only led to a better routine but helped the teachers trust me and the process.

Thus, all three participants found that the routine increased their perception of being treated as a professional during the teacher evaluation process.

**Teacher Evaluation**

Through the years, two competing purposes have been asserted for teacher evaluation—accountability and teacher development. Following World War II, a more developmental view of supervising teachers arose, which led to a great focus on classroom observation and clinical supervision. The clinical supervision model, attributed to Morris Cogan of Harvard University, was designed to focus on the development of the teacher with the teacher playing an active role in their own evaluation (Cogan, 1973). However, over time, Cogan’s systematic approach to teacher supervision turned into a more rigid process that betrayed its original intent (Marzano et al., 2011). On the heels of clinical supervision came Madeline Hunter, who introduced a structured seven-step framework for lesson design and instructional delivery that was often used in conjunction within a clinical supervision model. As with the unintended consequences of the structured clinical supervision model, Hunter’s framework was also often implemented rigidly.

Building upon Hunter’s idea of a common framework and language, Charlotte Danielson introduced her own framework, which broke down teaching into domains and elements and described different levels of proficiency (Danielson, 2007). Along with Hunter, Danielson and her framework are foundational to the standards-based approach to teacher evaluation that we currently use. In New Jersey, this move to standards-based teacher evaluation became the law with the passage of AchieveNJ in the second decade of the 21st century. The requirements of student achievement data and numerical rating pushed this initiative toward the accountability
end of the spectrum and came to dominate the discussion and therefore masked how the use of a standards-based teacher evaluation model could contribute to teacher development as an equal purpose of the teacher evaluation system. As stated in the previous section on teaching as a profession, there is tension between the existing system and the tenets of effective professional development. The existing system is more aptly designed to serve accountability than to serve teacher growth. The state attempted to accomplish the goal of promoting teacher growth through the use of a standards-based system, but the effects of including the student achievement data and a final numerical score were overwhelming and came to dominate.

The routine that was developed for this study focuses on teacher evaluation. Through this analysis, I believe that the routine worked well because it was characterized by a balance of standardization and flexibility that allowed for increased teacher agency, structured a narrow focus, and created collaboration between the teacher and observer. The teachers shared that they appreciated all these components of the new routine.

**Continuation of standards-based approach to teacher evaluation.** Our routine included clear, agreed-upon standards for teaching and common language, which the existing teacher evaluation system also includes. In gathering information from the participants about their experience with the existing system, I learned that they found the transparent standards and indicators to be the most beneficial aspects. This finding aligns with the literature that suggests a standards-based approach to teacher evaluation is more conducive to teacher development (Toch & Rothman, 2008). One way that we attempted to improve upon this was by narrowing the focus of the evaluation to a few agreed-upon indicators. The teachers found this approach more effective for improving their practice because it gave them a voice in choosing a focus for their observation and also made it more likely that they would improve a certain aspect of teaching.
Papay (2012) argues that we need to reframe teacher evaluation as a formative assessment tool with the goal of professional development. This reframing as a formative assessment is consistent with my findings that participants appreciated clear specific feedback and opportunities for further practice.

**Flexibility to use professional judgment.** One of the most important findings from the study was the importance of building flexibility into the routine. All the teachers believed that flexibility was important, because it both increased their perception of being treated as a professional and ultimately improved their teacher practice. As we sat down to revise the first routine in preparation for the second cycle, the teachers made the case for increasing flexibility and recommended that we avoid placing minimum requirements when possible. For example, when I suggested that we might consider setting a minimum number of lessons that each teacher must record and review before submitting, they argued that no minimum should be set, because such a criterion would sabotage the broader goal of designing a routine that increased teacher agency. So although the routine had a definite structure and procedures to follow, it was also designed so the teacher had great flexibility and choice. In connecting this to the history and literature surrounding teaching evaluation, I believe allowing for greater flexibility could help to guard against the rigid approach to teacher evaluation that resulted from Cogan’s clinical supervision model and Hunter’s framework (Marzano et al., 2011; Pajak, 2008; Wise et al., 1984) and the rigidity that resulted from the requirements of AchieveNJ.

**Effective Professional Development**

In Chapter 2, I identified the two main, and often competing, purposes of teacher evaluation: accountability and teacher development. Though not treated as a profession throughout its history, teaching is a complex endeavor that requires skill, expertise, and ongoing
decision making (Hawley and Valli, 1999). Therefore, teachers must develop their knowledge base and skills through practice-based learning and reflection upon their own practice through self-study and inquiry (Ball & Cohen, 1999).

It is within two conceptualizations of teacher development—practice-based professional learning (knowledge-in-practice) and teacher reflection (knowledge-of-practice)—that I examined the element of effective professional development. For this examination, I relied heavily on a recent review by Linda Darling-Hammond and colleagues (2017). In their review of 35 studies, Darling-Hammond et al. found that effective professional development consisted of the following characteristics: (a) provides feedback and reflection, (b) involves active learning, (c) is content focused, (d) includes coaching and support, (e) is collaborative, (f) uses modeling, and (g) is sustained. I used this definition of professional development to design my organizational routine that was the focus of this study.

As the designer of the original routine and a co-designer of the second routine, I paid close attention to the characteristics of effective professional development. In this section, I report what aspects of the routine the teachers found effective in improving their own practice, and, when possible, I connect these findings to the current literature on effective professional development as described above.

**Provides feedback and reflection.** All three participants found it extremely helpful and effective to have time-stamped comments providing specific feedback. Although they would receive feedback in the existing model, it would not be tied to a specific moment that can be reviewed using a recorded video. This level of specificity was important for the participants and in line with the research on effective professional development (Darling-Hammond et al., 2017). The participants found that this specificity improved the efficacy of the comments, because it did
not require them to remember the moment or decision that the observer was referring to. Furthermore, the participants reported that although the time-stamped comments were provided later, receiving them felt like receiving immediate feedback because the comments were tied directly to a moment the teachers could relive by rewatching the video while digesting the comment. In this way, time-stamped comments provided a type of practice-based professional development focused in the knowledge-in-practice conceptualization of teacher learning (Ball & Cohen, 1999). Moreover, the ability to attach links and suggestions to particular comments improved the level of support as reported by the participant, consistent with one of the characteristics of effective professional development (Darling-Hammond et al., 2017).

**Active learning.** Active learning, which enables teachers to learn in the same way they teach their students to learn (Drago-Severson, 2009), has been found important to effective professional development (Ball & Cohen, 1999; Hawley & Valli, 1999). The teachers in this study found themselves much more actively engaged in this routine than they were when participating in the existing teacher evaluation system. This active engagement came in a variety of forms, including the ability to record as many lessons as they wished, choosing the lesson for submission, completing the reflection, and watching scenes of the video collaboratively with me. Through these forms of active engagement, the teachers were positioned as researchers of their own practice developing knowledge rooted in their particular practice (Cochran-Smith & Lytle, 1999). Therefore, the new routine was situated within the knowledge-in-practice conception of teacher learning as opposed to the knowledge-for-practice conception, which defines the existing system. This reconceptualization fit nicely with the action research process we used, because the whole study aimed at developed knowledge from within the context of the classroom.
Another aspect that bears mentioning is the teachers’ responding to comments using the online platform before coming in for the postobservation conference. This became one of the most important aspects of the routine, because it positioned the teacher as equal participant in their own evaluation. It made the entire process more collaborative, including the postobservation, which was done in person. Going forward I would consider having the teacher comment and provide feedback on their own lesson before sending it to the observer for their feedback. While the teachers in this study did this informally, I would be interested to see how it would affect the experience to have them officially and formally document their self-critiques and reflections. This component was not included in the first routine, mostly because we did not know it was a feature of the software. Once we discovered it, I encouraged the teachers to respond to comments when appropriate in order to make the evaluation more of a two-way conversation and provide them with another opportunity to develop their knowledge-in-practice by using video. By asking them questions about decisions they made in the classroom, I prompted the teachers to place themselves back in the moment, thinking about their decision-making processes. If they had lacked the ability to dialogue with the outside observer, my evaluation would have been an exercise more aligned with knowledge-for-practice, in which the observer shares knowledge with the teacher, rather than an exercise in which observer and teacher coconstruct knowledge (Cochran-Smith & Lytle, 1999). The teachers appreciated this online discussion and also appreciated the ability to clarify a misconception or justify a teaching decision that was questioned. All in all, the inclusion of responses to comments increased the active participation of teachers and therefore increased the routine’s ability to function as a form of professional development.
A different form of collaboration. The participants found the routine we developed much more collaborative than the existing teacher evaluation system, because we collaborated as researchers, together creating, implementing, and assessing the new routine. One of the participants reported that this felt like something he was a part of rather than something being done to him. He attributed this feeling to several factors, including being actively involved in the revision of the routine, being part of this research study, completing a reflection before submitting the video, and watching parts of the video collaboratively with me as part of the postobservation. As part of the routine, the teachers had to complete a reflection prior to submitting their video. One of the findings was that teachers found this step important because it gave them more control over the process, allowed them to address issues proactively before receiving feedback from the observer, and overall made the postobservation more conversational. For example, Mark commented that he believed this component made the process more collaborative and leveled the playing field between teacher and supervisor. As discussed above in the section on teaching as a profession, the viewing of the video collaboratively during the postobservation also contributed to an increased sense of collaboration. (I discuss their participation in the study at greater length below.)

Although the literature focuses on teachers collaborating with each other as an important element of effective professional development (Darling-Hammond et al., 2017), we found that collaboration between teacher and administrator could also be perceived as having have a positive effect on teacher development, especially as part of the teacher evaluation process. As I reported in Chapter 4, this effectiveness of the collaboration was predicated on a strong relationship between the teacher and me. This trusting relationship evolved partly due to the design of the routine, which promoted such relational trust, but also because of certain actions
and approaches that I took as the co-researcher in this study. Although Darling-Hammond et al. (2017) leave open the possibility for many forms of collaboration, for the most part they are referring to collaboration between teachers. In this study, I found that collaboration between teacher and administrator can also be a very effective part of a teacher evaluation process designed to treat teachers as professionals and promote their learning.

**Contribution to Organizational Routine Theory**

As presented earlier, in Chapter 2, I used the idea of organizational routines—based in organizational theory—as the guiding framework for this dissertation. Specifically, I used Feldman and Pentland’s (2003) definition of *organizational routine*: “a repetitive, recognizable pattern of interdependent actions, involving multiple actors” (p. 96). Within the overall idea of organizational routine, it is important to differentiate between the ostensive routine and the performative routine (Feldman & Pentland, 2003). The ostensive routine is the routine in its idealized form, whereas the performative routine is the routine as it is actually practiced. When thought of within the context of teacher evaluation, the ostensive routine is how teacher evaluation is supposed to be carried out according to the design put forth in AchieveNJ along with specific district policies linked to that design (such as the selection of a specific evaluation instrument), whereas the performative is the act of teacher evaluation as it is locally implemented. In this section, I begin by revisiting the value of routines in education and then attempt to discern which of my findings can be attributed to the ostensive routine and which findings are more appropriately attributed to the performative routine.

**Revisiting the Value of Routines**
As described in the literature review, organizational routines have the power to create substantive changes in educational environments (Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). These changes have been seen not only in mundane aspects of schools but also as a catalyst for changing of instructional practices that are at the core of teaching. However, changes to organizational routines are often resisted (Akiba & Wilkinson, 2016; Conley & Enomoto, 2009; Sherer & Spillane, 2011; Spillane, Parise, & Sherer, 2011). Therefore, it is important to be mindful of teacher agency issues when one is designing any change to an existing routine or creating a new one. Furthermore, studies have found that teachers can derail any change, because they are the ones who will need to implement it. Therefore, the studies reviewed for this dissertation suggest that a school leader continually get feedback, reassess the routine, and address resistance to the routine. With careful design and stewardship, an organizational routine is more likely to succeed. And, if successful, organizational routines can both change practice and then help that change become formalized as a new constant in the organization (Sherer & Spillane, 2011).

In designing and carrying out the organizational routine for this study, I found that organizational routines can indeed change teacher perceptions. Moreover, I found that the positive changes were sometimes attributable to the design of the ostensive routine, whereas others are more appropriately traced back the performative aspects of the routine. In terms of my study, I found that trust is paramount for developing a new routine—especially one dealing with an issue that can be anxiety producing, like teacher evaluation. As I presented in Chapter 4, trust was developed in several ways with the group in this study. In the literature, researchers recommend that school leaders get feedback on their changes to help improve the routine and to increase the probability that it will be accepted by the teachers and staff who will need to enact
it. In my study, another powerful model is positioning the teachers as co-designers of the routine, rather than consultants giving feedback. By framing the teachers as co-designers and being transparent about the purpose of the routine, I made it easier for the teachers to invest themselves in developing and improving the routine for their own benefit and for the benefit of teachers at large. As the administrator and leader researcher, I intentionally gave some control of designing the routine to the teachers. As I was able to this, the teachers appreciated that I was putting into practice what I stated was the purpose going into the study. This alignment of words and actions further developed trust between me and the teachers.

The Ostensive Routine

The ostensive routine is the idealized form of the routine as it is conceptualized. In this study, I employed two versions of an ostensive routine—a version designed by me and enacted in the spring, and a second version co-designed by me and the three teachers and carried out during the fall. The first version was based on the literature reviewed in Chapter 2. In the design, I included lessons from the literature on effective professional development and organizational routines in education. In the second version, I kept in mind the findings of the literature while also giving credence to the teachers’ experience during the first routine. In each of the examples below, I briefly show how a component was designed intentionally and how it was experienced by one or more participants while it was enacted.

I set out to design a teacher evaluation routine that met several of the characteristics of effective professional development. With that in mind, I specifically included components that promoted active learning that combined feedback to the teacher along with opportunities for the teacher to reflect on their practice. The process of choosing one video for submission after an unlimited number of recordings was one component that I designed with these characteristics in
mind. Giving the teachers the opportunity to record and watch their own lessons before deciding upon one for submission afforded them an active role in the process. They were the ones who had control of the recording device. They had control of which lessons they chose to record. They had control over which lesson they ultimately chose for submission. Once they narrowed their lessons down to the one that they wanted to submit, they were required to complete a reflection that further encouraged active learning on their part. The participants found that this process did indeed engage them in active learning and that much of their professional learning took place before they even submitted a video for evaluation. Therefore, in this particular case, the routine was enacted as designed. As we moved from the first to the second version of the routine, the teachers wanted to keep these components, with one adjustment being they had an even longer window of time in which to record their lesson and choose their final lesson for submission.

Another part of the routine that was designed around effective professional development and that ended up being perceived by the teachers as a positive component was the time-stamped, specific feedback aligned to the Stronge teacher performance standards. As I designed the routine, I wanted to leverage the power of video and the online platform for providing feedback directly tied to a moment so that the teachers would ostensibly get richer and more detailed feedback about their practice. The participants found this to be a valuable component, and, for the most part, they implemented it as planned. Two ways that the feedback feature was enhanced arose organically as we performed the routine together during the first cycle. One was that I provided links and attached resources to specific comments, and the second was that in each case the teacher responded to specific comments. Both of these enhancements were
incorporated into the second cycle and became codified as part of the second version of the routine.

Based on the findings from literature on organizational routines in education, I was aware that paying attention to teacher agency was important for implementing an organizational routine (Conley and Enomoto, 2009). Attending to teacher agency is also important whenever one wants to effect a change in order to increase a teacher’s sense of professionalism, and therefore the concept of teacher agency became doubly important for my study. As a result, I framed my study so that the teachers would serve as coconstructors of the second version of the routine. Based on their experiences with the first version, the teachers knew going in that they would have the opportunity, and obligation, to suggest revisions for the second version. This work resulted in five changes to the first version that enhanced alignment with effective professional development. For the purpose of review, these five changes were a longer time window for the teachers to film their lesson(s), an awareness of and emphasis on teachers using the comment feature in Swivl Cloud to respond to comments, a more concise reflection, a focus on fewer scenes during the collaborative viewing, and an optional reflection or submission of artifact following the observation cycle. All of these changes were made for the purpose of increasing teacher agency in the process or providing more targeted feedback to teachers on their practice.

**The Performative Routine**

As shown above, the ostensive routine was designed to promote a sense of professionalism and foster teacher growth. Many parts of the routine were enacted as designed, and the positive findings could be fairly attributed to those components. However, there was also the performative routine, which was not always enacted in the way the routine was designed. Furthermore, the relational aspect of the routine and process should not be
overlooked. As a group, we implemented this routine in a certain way and in a certain context. This included specific moves by me that fostered trust between participants. For example, I explicitly stated the purpose of the study, and therefore the teachers knew going in that this routine was part of the study for my dissertation and might be implemented across the district in coming years. Both of those factors contributed to the increased perceived importance of the process and of their role in it. Although certain outcomes of the study can be traced to the design of the routine as discussed above, the process of performing the routine also contributed to certain findings. For example, the teachers shared that they believed that participating in the research and the co-design of the second version of the routine affected how they experienced the routine. They remarked that they appreciated that fact that they were working on something that could ultimately be used in the district. They felt that they were part of something important. They also remarked that some of the procedural aspects of the research study contributed to their sense of belonging and membership in a special group. For example, they referenced how trust was built through frequent email communications and several visits to my office for interviews. Although these experiences were not explicitly intended to foster trust, it is interesting to note that they did. In thinking about how to formalize this, an administrator might think about conducting action research with a small group of teachers and making one goal of the research be to continually consider revising the process and procedures of the teacher evaluation routine. This would project the belief that the evaluation routine is not fixed, but always revisable, and that both teachers and administrators have important perspectives to add to this iterative process. However, I am not sure that this type of collaborative research is doable at scale across a district, because it is time intensive and part of its power was the fact that the
administrator was working with a small group of teachers. If this were attempted at full scale across a district, I fear it would lose its authenticity as a truly valid way to revise a routine.

When asked if this process caused her to see teacher evaluation differently, Elise shared that it did, because she participated actively both in the designing of the routine and in the evaluation of her own teaching. This feeling of playing a more active role in their own evaluations was something that all the teachers described experiencing. They felt like collaborative partners who had real agency and control in their own evaluations. Evaluation was something being done with them and not to them.

As was shared earlier, Jonathan’s department was very skeptical about the entire process when he shared with them that he would be participating in this study. At our final group interview, unsolicited, Jonathan shared that he had come to see real benefits of the routine and found that it could increase teacher agency, control, and professionalism rather than hinder it, and he was sure that “if teachers really understood what the use of this was for, they would jump on board” (Jonathan, Group Interview 2, October 30, 2019). This skepticism of Jonathan’s colleagues is important to note and be mindful of when looking to scale this routine up across the district.

As I described in Chapter 4, the teachers shared that they felt very comfortable trusting me as the administrator who was working with them to carry out this new routine. As I looked back upon the process and reflected upon the teachers’ comments that they strongly trusted me, I thought it would be informative to self-assess and share what I did to build this trust. First, I believe that my reputation in the district positioned me to go into the process with a certain degree of trust already earned from the participants. I have worked with all of them in some capacity, and throughout those interactions I believe we have developed mutual trust. Before
embarking on this research project, I did work more closely with some than others, though, and I believe the level of trust cultivated can also be attributed to other factors. One of these factors is how transparent I was with the teachers throughout the process. I shared the purpose of their participation, how I thought they would benefit, how I would benefit from their participation, and how I truly wanted to partner with them to create a scalable new routine for the district. The transparency was also coupled with constant communication via email and in-person visits. I tried to make participation in the process as seamless as possible and always offered to work around their schedules and meet at times and locations convenient for them. However, it is important to acknowledge that all three teachers volunteered to be part of this study, so it can be inferred that their experience might not be representative of teachers in general.

As described in the literature on organizational routines, changes to routines are often resisted, and leaders would be wise to be mindful of teacher agency when designing and implementing any change to routine. I believe my study shows that a leader can help increase the success of implementing a new routine by having a group of teachers help codesign said routine. Finally, I believe a lot of the trust we had for each other in the second cycle grew out of our collaboration during the first cycle, in the spring. The teachers found value in the process, appreciated my thoughtful and specific comments, and were grateful that we did come to coconstruct the second version of the routine for the fall cycle, as I had promised at the outset. In working closely with the teachers in performing this routine on a small scale, I encountered several questions about how this could look when scaled up across the district. In an upcoming section on recommendations for district leaders, I refer back to these discussions when making specific recommendations for scaling up. In the final reflection, I return to some of these lessons learned as a leader using action research to introduce a change in practice.
Recommendations and Implications for Future Research

In this section of the final chapter, I share my recommendations and identify some implications for teacher evaluation at both the district and state level. Despite the positive findings of this study, it is important to note that one faces real tension when attempting to design a teacher evaluation routine that fosters teacher learning, especially on a large scale. Nevertheless, in this study, all three participants found that the organizational routine we developed enhanced their perception of being treated as a professional and served well as a tool for improving their practice. Based on all our experiences across the two cycles of this action research study, I have several recommendations for how the findings of this study could improve existing teacher evaluation systems. My recommendations are aimed at two separate groups: leaders in any school or district, including my own, and state policy makers in New Jersey. Because one of the stated goals was to develop an organizational routine that could be scaled up in my district, I pay particular attention to my context for the recommendations.

Recommendations for District and School Leaders

In this section, I share my experience using action research as an educational leader while considering my next steps, before sharing some implications for other leaders who might engage in this type of action research with teachers moving forward. Participating in this action research study turned out to be a wonderfully informative and educational experience for me. It provided me a chance to collaborate with teachers from three of the five schools in the district as I set out to document both the teachers’ experiences and mine while we examined the creation of a new organizational routine that could be implemented in our district in the coming years.
Practical changes to routine when scaling up. Based upon the success of this study, our district will be looking to scale up this approach to teacher evaluation. I also hope that leaders in other school districts will examine the findings of this study and decide to implement a similar routine in their district. As we prepare to do this, we must consider some pragmatic factors in order to give the expansion the highest probability of success.

Keep it optional. First, I highly recommend that this approach remain optional for teachers and not be mandated for any of them. One of the most powerful aspects of the routine was the fact that the teachers themselves chose to participate, which immediately increased their feelings of teacher agency and buy-in to the process. I believe that mandated participation would kill that aspect of the routine, which in turn could sabotage the entire routine.

Give teachers a large window of time. Second, I would recommend continuing to offer the larger, five-week window for teachers to film as many lessons as they want and then choose one lesson for submission. The two-week window of the first cycle was rushed. The windows could differ for each teacher involved, and a district can build in teacher choice for what time of the year they would like to conduct their observation cycle. This would not be a hard change to make and is not much different from the current system used in our district. Currently there are observation windows to help balance observation loads for administrators and spread out observations for teachers. In this new paradigm, we could still use the window concept, but now the teachers would have control over when they conduct their lesson and submit their video.

Streamline the number of online platforms. A third change that I would make would be to look for more efficiencies with the process to streamline it for both teacher and observer. For example, in the first and second versions of the routine the teachers were required to complete a Google Forms reflection before submitting their video. Based on Elise’s suggestion, I would like
to build these reflection questions directly into Swivl Cloud so that the teachers would not have to toggle between the two platforms. From an administrator perspective, having to translate the feedback from Swivl Cloud to our district’s evaluation platform, Frontline Professional Growth, was a little cumbersome and added an extra step to the process. I recommend researching a way to eliminate this step.

Guard against blurring of boundaries. One of the best parts of this routine for an administrator is the flexibility that it offers. As shared earlier, the use of video allowed me to conduct observations while in my office, at home while my daughter napped, and really anywhere with an internet connection. Although this flexibility is terrific, an administrator must be mindful of the positive negative effects of this constant connectedness and the blurring of work and home life. Administrators are currently dealing with issues of blurred lines between home and work, because they are always connected and on the job. As a former building principal, I lived this reality for many years and am acutely aware of how the job can encroach on one’s personal life. Any administrator who is going to participate in a version of this routine should go into it with the awareness of this possible pitfall and consciously guard against it by setting parameters and guidelines for themselves.

Substantive changes to routine when scaling up. In this section, I recommend steps for easing or resolving the tension between teacher evaluation and teacher learning. Because these are recommendations for the district level, they all take into consideration the current state requirements, and therefore all recommendations made satisfy those requirements.

Skill and stance of observer matters. When discussing trust in Chapter 4, I presented data that showed that the teachers were aware of my stance as a researcher, trusted my motives, and also had confidence in my skill as an observer of teaching practice. The teachers appreciated my
stance as a researcher who did not come into the collaborative relationship with the answer already in mind, but wanted to examine this routine and my own practice as it related to that routine. As the observer, I found that my skills as an observer were paramount to the success of the routine. In other words, the routine is not designed to be “administrator proof.” It does not relieve the administrator of using their judgment, and it requires them to observe thoughtfully, provide specific feedback tied to standards of practice, offer suggestions, and provide support where appropriate. Therefore, I recommend the development of an administrator training program that focuses on how to use video to enhance elements of effective supervision.

*Relationships matter.* The other part of the process that cannot be overlooked is the relationship between teacher and observer. As was shown in the findings, the teachers in this study shared their high level of trust in me and my abilities as an observer. This trust built up partially due to their role in this study as coconstructors of the second version of the routine. If the routine were scaled up to more teachers in the coming years, those later teachers would not have played the same role, so it is imperative that administrators pay attention to the relational aspect of the routine. They must guard against the use of video and an online platform becoming a barrier between teacher and observer. Furthermore, the administrator involved must take the privacy issues of the use of video with the utmost seriousness. As I discussed, many of Jonathan’s colleagues were very skeptical of the uses of video due to its permanency and the fact that the teachers are sending something out into the universe that can be shared with others. The teachers must trust that the administrator will not share the video of their teaching without their explicit permission. To do so would betray the trust that the teacher placed in the administrator and the process and would most likely doom the initiative’s success if word spread that the videos were being shared without permission.
Purpose matters. An important finding of this study was that although the design of the routine was important, equally important how the observer performed it. As stated earlier, the routine is not administrator-proof and is not a step-by-step procedure that can be followed blindly. The observer needs to pay attention to issues of trust as well as develop skills in offering helpful, targeted feedback in this new medium. I found that this routine can be successful when implemented under the correct conditions. One of the most important conditions is the approach that the administrator takes in implementing the routine. When I set out to undertake this study, I was curious if the routine would save time. In our implementation across two cycles, I did not find that it did. If an administrator wanted to use this routine to save time, they definitely could, but the time savings would almost assuredly come at the expense of a high-quality observation focused on the development of the teacher in question. For example, an observer could quickly go into a video file, skim through, and add a few nonspecific comments on various parts of the video. Although this would save them the time that usually would be spent in the classroom, it would not accomplish the larger goal of using this routine to reframe observation as an effective method of professional development.

Recommendations for New Jersey State Policy Makers

Expand the pool of candidates. The first recommendation has to do with the pool of teachers who can participate in this type of differentiated supervision that increases teacher agency. Currently in the state of New Jersey, only teachers who have been rated as highly effective are permitted to participate in the differentiated supervision that emphasizes teacher reflection and includes a video component. It is interesting that none of the three teachers who participated in my study would have qualified using the above-mentioned criterion. All three were rated effective using the district’s teacher evaluation system the previous year. In two
cases, this was due to the weight of the student achievement scores in the teachers’ final numerical rating. Nevertheless, as I have shared in the findings chapter, all three participants believed that they benefited from participating in this process and that it increased teacher agency, treated them as professionals, and provided more specific and targeted feedback.

**Increase flexibility.** In my opinion, the existing model used by the state of New Jersey, allowing differentiated supervision only for highly effective teachers, is not flexible enough. Besides the videotaped lesson, it requires other components—such as conducting a student focus group and keeping a reflection journal—that could make it less attractive to a teacher. The teachers in my study thought that a weakness of the current system was the amount of time and effort the paperwork required. Therefore, leaders might want to identify components to remove or replace and limit the addition of new requirements. The current option for highly effective teachers also requires that the videotaped observation replace the announced rather than the unannounced observation. The participants in my study believed that the videotaped observation would have much more potential as a replacement of the unannounced observation. They considered the unannounced observation counterproductive to their development as teachers, and they still appreciated participating in one full evaluation based on the clinical supervision model consisting of a preobservation conference, and in-person observation, and a postobservation conference.

As we have discussed, teacher evaluation and professional learning have always been in tension. Although some teacher evaluation programs have been designed to address teacher learning, these programs often became bureaucratic exercises in accountability that, in practice, lose all of the components that were originally included to enhance teacher learning. In the design of the routine for this study, components were built in to increase the flexibility teachers
had to make decisions. For example, they were given a large window during which to film their lesson and choose which video to submit. Within that window, the procedure for recording and choosing was left pretty open, with much of the decision making falling on the teacher. This was meant to increase teacher agency, and it ended up doing that. Each of three participants ended up choosing a slightly different approach and, over the course of two cycles, ended up filming and watching more videos than they said they would have if a strict process had been set, with a minimum threshold. Therefore, I recommend that policy makers not set a minimum number of videos to record or watch and that overall this process be kept fairly flexible to allow the teachers to use their discretion on how the process will best work as a professional learning opportunity for them.

**Implications for Future Research**

The action research outlined in this dissertation took place over two cycles. After the first cycle the participants and I reflected and revised the organizational routine for the second cycle. Although the dissertation research “ended” at the end of the second cycle, I will continue to move this research forward into another cycle as I try scale a version of this routine across my district. Therefore, the research is in no way complete. For my particular context, I will be opening up participation in this alternative form of observation to a larger group of teachers. In the current research, the group was fairly small, with only three teachers. I believe future research should include a larger group of teachers, which would help diversify the findings and show whether the findings of this dissertation are reflected in the experiences of other participants, who may not be as deeply involved in the research itself. Along with including a larger group of teachers in the next phase of research, I would include other administrators besides myself. Although my experience was worth examining and reporting out, other
administrators—especially building principals—would experience it differently, and their experiences would be worth learning about.

Outside of my district’s context, I believe future research could focus on how the use of video for teacher evaluation would be experienced in different settings and with a different teacher pool. This study took place in an upper-middle-class, fairly homogenous school district with a history of teacher leadership and collaboration between teachers and administration. Future studies should focus on action research around how this type of organizational routine would be experienced in schools and districts with more diversity and lower socioeconomics. Possible areas to focus on could be trust between teachers and administrator, professionalism, and teacher agency. Furthermore, this study included three participants who were all tenured and were either in mid-career or late in their career. Future research could focus on how this routine is experienced by nontenured teachers with fewer years of experience.

Other possibilities for future research could be examining the possible value of organizational routines that are developed via action research that involves collaboration between administrators and teachers. The power of organizational routines is their ability to bring about change and then sustain that change. In this study, I focused on an organizational routine dealing with teacher evaluation. Future studies could focus on a myriad of different organizational routines that occur in schools.

Finally, future research could focus on other ways that video could be used for teacher learning. Although this dissertation focused on using video for teacher evaluation, video can support teacher learning in other ways as well. Future research could focus on some of these ways, including using video for self-reflection and peer collaboration. For example, it would be interesting to study the design of a new organizational routine around teacher evaluation that
incorporated self-reflection or peer observation in place of the administrator conducting the observation. This study opens the door for a wide array of future scholarship that incorporates the use of video, organizational routines, action research, and teacher agency in any number of educational improvement efforts.

**Final Reflection**

Teachers have had a hard time earning their place as true professionals, and this difficulty has led to a transmission approach to teacher learning and a managerial approach to teacher evaluation. Through the codesign of a new organizational routine we set out to use video to reframe teacher evaluation as a form of effective professional development. In doing so we hoped to increase both teachers’ perceptions of being treated as a professional and their perception of the routine as a way to help them improve practice. In documenting the teachers’ experience and mine, I found that this routine shows promise as an approach that could accomplish these two goals. Moreover, I found that the process of working closely together on the design of the routine seemed to have a positive effect on the teachers’ experience with that routine and mitigated resistance that might be expected with the introduction of a new routine.

Through participation in this qualitative action research study, I learned multiple lessons that I will take with me as I continue to try to promote positive changes as an educational leader. First and foremost, I learned that qualitative action research can be a powerful catalyst for change. It provides the opportunity for a school leader to work collaboratively with teachers on a problem of practice that is relevant to all parties. Specifically, I learned that clearly and honestly describing the process and stating its purpose at the outset of the study helps the leader win teachers’ trust and keep the process focused on the core purpose of the research. Furthermore, I learned that a leader must be truly open to feedback from the teachers and not
have a predetermined outcome in mind when beginning a study. However, I also believe that this study succeeded partly because I worked hard on designing the original routine and attended to all the details of the process. The teachers appreciated having a framework to start with, rather than having to design the entire routine from scratch, and also appreciated that they were given a clear process with support for them built in along the way. This is key for administrators who want introduce a change to an organizational routine, conduct action research, or, in fact, bring about any change to a school organization. Be clear about the purpose, be open to coconstruction with your stakeholders, and do not go into the process with the result predetermined. The process is messy, in that you are not in control of how things will play out, and this requires a tolerance of giving up power. In my experience, this is difficult for many administrators to do, myself included. Going forward, I will be looking for ways to work more collaboratively with teachers when I hope to bring about change. In fact, I have already started.

In conclusion, I hope that the findings of this dissertation help scholars studying organizational routines by providing a rich example of how a transparent, trusting, collaboration between teachers and administrator can help in the design and implementation of a new routine. I also hope that these findings prove helpful (a) to my district when it tries to scale up this routine, (b) to any district hoping to introduce a similar routine, and (c) to any educational leader looking to partner with teachers through action research. Finally, on the state level, I hope that the findings of my research help policy makers from the state of New Jersey work collaboratively with educational leaders and classroom teachers to codesign a new evaluation routine for teachers that includes the use of video, is open to a larger group of teachers, remains optional, and consists of real choices for the teacher involved in their own professional learning through evaluation process.
REFERENCES


doi:10.1080/02660830.2007.11661545


doi:10.3102/0002831216630407


doi:10.1177/003804070407700302


doi:10.1177/0022487108328155


through videotaping and guided assessment. *Journal of In-Service Education, 34*(1), 75–
95. doi:10.1080/13674580701828229

mechanisms: Policy, school administration, and the technical core. *American Educational

evidence from a teacher evaluation pilot in Chicago. *EducationNext, 15*(1). Available at
http://educationnext.org/better-observation-make-better-teachers/

Supervision and Curriculum Development.

Stronge and Associates (2016a). *Frequently asked questions about the Stronge Evaluation


Stronge and Associates (2018). *Does teacher evaluation improve schools and student learning?
What does the research say?* Williamsburg, VA: Strong and Associates.


https://www.rand.org/pubs/reports/R3139.html


doi:10.1177/0013161X15616661

APPENDIX A

Reflection Questions

1. What was it like watching your own classroom lessons?
2. What did you notice?
4. Did you see anything in the videos (students’ behavior, things you thought you said, things that you did not remember saying) which you were not aware of while you were teaching the lesson?
5. What was your area of focus for this lesson?
6. What did you notice in your area of focus while watching your lesson?
7. What type of supports or professional learning would help you grow in this area of focus?
8. After viewing the video, what would you consider the next steps for your practice?
9. Were there other areas that you noticed that you would like to consider addressing?
10. Is there anything else you would like to share about the experience?
This page intentionally left blank