WORKING ON COHERENCE WHEREVER THE TWAIN SHALL MEET
HOW TEACHERS CONTRIBUTE TO POLICY COHERENCE THROUGH ADAPTIVE RESPONSES TO RELATED DEMANDS OF EVALUATION AND PROFESSIONAL LEARNING COMMUNITY

By

Dirk Frederick Zuschlag

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Educational Policy – Doctor of Philosophy

2021

ABSTRACT

WORKING ON COHERENCE WHEREVER THE TWAIN SHALL MEET
HOW TEACHERS CONTRIBUTE TO POLICY COHERENCE
THROUGH ADAPTIVE RESPONSES TO RELATED DEMANDS
OF EVALUATION AND PROFESSIONAL LEARNING COMMUNITY

By

Dirk Frederick Zuschlag

Teacher quality is a perennially fertile field for education reformers. Among teacher quality policies, two of the more highly touted and widely adopted are state-mandated systems of teacher evaluation and formalized models of teacher collaborative teams, often known generally as professional learning communities (PLCs). Indeed, despite a substantial divergence in policy design and implementation process, many districts and schools have by now routinized the annual re-enactment of evaluation and PLCs policies, thereby engendering a recurring challenge of "crafting coherence" (Honig and Hatch, 2004). Teachers in particular, both individually and in PLCs, must fashion responses to disparate demands, which interact to affect but not determine those responses. How teachers respond holds important implications, not only for the iterative implementation and effects of potentially incoherent policies such as PLC and evaluation, but also for larger teacher and school improvement efforts and outcomes.

This explanatory multiple case study in two public high schools investigates teacher agentic responses to dual teacher quality demands within distinct but related PLC/evaluation structures applying a structure-agency perspective (Coburn, 2016). It addresses the following research questions: (1) How, if at all, do teachers adapt their agentic responses to performance evaluation and professional learning community demands as they respond to the structures of both; and (2) how, if at all, does the relationship between structures of performance evaluation and professional learning community shape teachers' adaptive responses?

Results show that teachers generally compartmentalize their respective responses to PLC and evaluation demands. Importantly, however, periods of compartmentalization are punctuated at points in the evaluation process. Punctuation occurs when structural opportunities open for teachers to advantageously adapt their response to one demand to serve their response to the other demand. Although teachers vary in the timing and form of punctuation, four identifiable types of adaptive responses emerge from teacher agentic action within the related structures administrators implement. A typology of these strategies is proposed based on the placement of each along two intersecting dimensions. When teachers employ the strategies, they can reduce the cost of evaluation engagement, while increasing the value of PLC participation. At the same time, teachers were in effect able to enhance the coherence of their responses.

Further results show that related PLC/evaluation structures are characterized by a design orientation--primarily commitment or control (Rowan, 1990)—and a goal orientation--primarily external or internal relative to the PLC/evaluation structures. One case high school implemented related structures with control-external orientations, the other commitment-internal. It is these paired orientations as implemented that can significantly influence when and how teachers may use certain adaptive response strategies.

The implications of these findings are discussed, including those that involve the application of structure-agency theory in education policy research, the understanding of how coherence may be crafted in routine, multiple policy implementation at the "street level," and for the work of policy makers, practitioners, and researchers.

I dedicate this dissertation to my wife Sharon.
Almost four decades ago, you chanced that “we’d make a great team.”
And so we have, through all the turns in my vocational path, among
all the vagaries of our wonderful life together, because of you. Thank you.

ACKNOWLEDGMENTS

No endeavor like this is ever one person's product. This work being no exception, I owe many thanks to many. So many, in fact, that I fear inadvertent omissions. Nevertheless, I want to try, poorly perhaps, to express my gratitude to some of those who accompanied me in process and contributed to its result through unwavering care, help and support.

I first want to thank the educator participants of the pseudo-anonymous Woodford and Walker districts and high schools, the sine qua non of my research. These consummate professionals, and especially the focus teachers, gave thoughtfully, openly and generously of their time, experience and knowledge. My deep appreciation for your dedication to what you do only grew during the course of the study.

In a similarly direct connection with this research, I owe much to my dissertation committee chair, Rebecca Jacobsen, and committee members, Ken Frank, Kristine Bowman, and Chris Torres. Scholars *and* educators *and* kind people all, I thank you for your consideration and encouragement, expectations and challenges. Further, beyond committee service, beyond impactful coursework, each committee member in their own way played a mentoring role in my ongoing professional development. Dr. Jacobsen, my advisor from day one, was constantly engaging me with the practicalities, the realities of research, writing, and the world of academia. I have had many good experiences and grown in many ways from my participation in Dr. Frank's Balancing Voices project and with his research group, as well as from my work as Dean Bowman's research assistant. And, even before joining my committee, Dr. Torres helped me think more clearly about my research interest and approach. Again, thanks to each and all.

A number of other MSU faculty members and researchers have similarly helped me on the way. I would like to thank Dr. Michael Sedlak for encouraging me to enter the program, for leading the first year ED 928 proseminar, and for providing support throughout the program. Thank you, Dr. Sarah Reckhow, for introducing me to your work and methods, assisting me with my coursework in political science, and serving on my guidance committee. I am indebted to Drs. Nathan Burroughs and Jacqueline Gardner for bringing me into the substitute teacher research team and providing for my meaningful involvement in practically all phases of the project. I also continue to learn from Drs. John Lane and Kaitlin Torphy Knake, my far more knowledgeable and experienced colleagues in Dr. Frank's Balancing Voices project.

I would like to thank my fellow ed policy graduate students, who were to a person welcoming, considerate, and generous to one from a decidedly distant generation. I derived particular benefit from the practical counsel, example, and encouragement of graduate student leaders, such as Ben Creed, Rachel White, David Reid, Jason Burns, and Kim Jansen Evert. Many thanks. Thank you, too, to the students participating in Dr. Frank's research group.

Finally, I am exceedingly grateful to my wife, Sharon, to whom this work is dedicated. I am regularly astonished by, and so absurdly fortunate for, her ability with good cheer to provide for all facets of our well-being. I also thank family and friends for their understanding, interest, and encouragement.

I am and will remain ever thankful both for those recognized above and for many others who, but for the finite nature of space and time, I would likewise acknowledge for furthering my work in education, culminating in this dissertation. Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Over more than three decades, an education reform wave has inundated K-12 public education at every level. The torrent of policies has been, and remains, characterized by top-down state and federal mandates with shifting policy fields, targets and mechanisms. Even as they scramble to handle external demands, however, districts and schools still endeavor to implement internal improvement goals and strategies (Honig, and Hatch, 2004). The result, as Honig and Hatch wrote in 2004, is "a heightened state of policy incoherence" at the local level (p. 17 [references omitted]; see Fuhrman, 2003a, 2003b; Newmann, Smith, Allensworth, and Bryk, 2001). Whether externally imposed or locally adopted, the initial implementation and eventual institutionalization of incoherent reforms increases the likelihood of policy and resource conflict, waste and dysfunction, as each reform undermines or impairs the others (Robinson, Bendikson, McNaughton, Wilson, and Zhu, 2017; Srinimsan and Archer, 2018; Wood, 2007).

> Multiple simultaneous school reform initiatives, each with their own accountabilities, can fragment the efforts of leaders and teachers, create stress, and increase cynicism without achieving the intended improvement (Hess, 1999).

(Louis and Robinson, 2012, p. 633). Of course, the casualties of policy incoherence extend beyond school improvement efforts to, ultimately, student learning.

One perennially fertile field for education reformers is teacher quality. Contemporary reform gospel, after all, holds that teachers make the single largest impact on student achievement (e.g., Lankford, Loeb, and Wyckoff, 2002). Unfortunately for teachers, the corollary often follows that teachers are principally responsible when student learning fails to improve. Among teacher quality policies, two of the more highly touted and widely adopted are state-mandated systems of high-stakes teacher performance evaluation and formalized models of

teacher collaborative teams, often known generally as professional learning communities (PLCs). These differ greatly in policy framing and design, implementation process and function. Still, the two in practice can overlap and interact through similarities in goal (e.g., teacher improvement through professional learning), subject matter (e.g., curriculum standards, instructional strategies, aligned assessments, student achievement data), and teacher activities and artifacts (e.g., lesson planning, assessment forms, student learning measurements). Many districts and schools have by now institutionalized and routinized PLCs and evaluation so that both are together re-enacted each school year.

The demands of these two divergent teacher quality policies annually converge and interact on teachers, who must of necessity respond to both (Knapp, Bamburg, Ferguson, and Hill, 1998). Like the reform initiatives of concern to researchers such as Knapp, et al. (1998), PLC and evaluation cannot help but "separately and jointly alter, refocus, enhance, complicate, or otherwise affect" teachers' work, working conditions, indeed their working lives (Knapp, et al., 1998, pp. 398-99). Convergent reform demands may disrupt, divert or constrain educator capacity, individual and collective, to improve practice through the development of teacher professionalism and professional accountability. In any given context, the result may not be for the good of teachers or students because

> …well-intentioned initiatives wind up creating conflict for practitioners and those they serve. Different initiatives may clash with one another, with the contexts in which they are implemented, or with the realities of how people respond to and adapt to change.

(Srinivasan and Archer, 2018, p. 3). Furthermore, whereas district and school administrators may and should act to meet the coherence challenge resulting from policy convergence (Honig and Hatch, 2004), teachers, the ground-level policy targets directly responsible to their students and parents, must somehow navigate the policy interactions and effects. How teachers respond, and

why, holds important implications, not only for the iterative implementation and effects of policies such as PLC and evaluation, but also for larger school improvement efforts and outcomes.

This embedded, explanatory multiple case study in two comprehensive public high schools applies a structure-agency theoretical perspective (Coburn, 2016; Donaldson and Woulfin, 2018; Rigby, Woulfin, and März, 2016) to focus on teacher agentic responses to dual teacher quality demands within distinct but related evaluation/grade-level PLC structures. More specifically, this research investigates how teachers adapt their response to the demand of one structure, evaluation or PLC, through their response to the other. It in addition shows how such teacher adaptive responses are shaped by any interaction between evaluation and PLC structures as administratively implemented. There is a pas de deux between teacher agentic responses that is influenced by, but that also influences, a pas de deux between structures.

Overall, I find that the predominant, day-to-day state of teacher responses to have involved teachers compartmentalizing their respective responses to PLC and evaluation demands. However, these periods were punctuated at certain points during the evaluation process when structural interactions of demands opened opportunities for teachers to advantageously interact their responses such that the response to one demand could be re-purposed, leveraged, or otherwise adapted to serve the response to the other demand. Although teachers varied in the timing and form of the opportunities, and therefore in how they adapted and applied their responses, four distinct types of adaptive response strategies emerged from teachers' agentic action within the related administrative structures. Teachers' motivation to engage in "punctuated compartmentalization" stemmed from a combination of their main goals in responding to evaluation and PLC, and how they valued their engagement with each. These

considerations also help explain why it was an evaluation demand that almost always triggered teachers' use of an adaptive response strategy. Once prompted, teachers took PLC work done to meet PLC demands while pursuing PLC goals and applied it in response to evaluation demands while pursuing evaluation goals. From a teacher's perspective, adaptive response strategies available due to collective PLC work at once served to reduce the cost of evaluation compliance and to increase the value of PLC engagement. And in the bargain, teachers were in effect able to enhance the coherence of their responses.

If teachers engage in a form of coherence crafting by how they adapt their agentic responses, administrators may facilitate or impede coherence by how they establish and implement the relationship between evaluation and PLC structures. In light of their larger policy attitudes, experiences and objectives, the administrations under study took differing approaches to their respective structural relationships based upon how each perceived the means to attain coherence while still achieving their preferred teacher quality improvement goals. Variation in structural relationships leads to variation in teacher response because certain structural characteristics can significantly affect the interactions that provide the opportunities for and teachers use of different adaptive response strategies.

For impact on the scope and nature of teacher adaptive responses, the administratively determined and implemented structural relationships varied most significantly in their design and goal orientations. Thus, I find that the alternative structural relationships, first, were principally control oriented or commitment oriented (Rowan, 1990; see Darling-Hammond, Wise, and Pease, 1983 [rationalistic versus natural systems model]; Ingersoll, 2003 [bureaucratic versus professional model]). Second, the administrative goal animating the structural relationship was externally or internally goal oriented, where an external goal is one over and above, or extrinsic

to, the goals of PLC or evaluation per se, whereas an internal goal is one intrinsic to or bound up with the achievement of PLC or evaluation results. Whether external or internal, the PLC/evaluation relationship is a structural means to an end. However, in the external goal orientation case, the end lies outside the intended purposes of both evaluation and PLC themselves; in the opposite case, the end lies inside their intended purposes—i.e., it *is* an aim of PLC and/or evaluation to begin with. In one of the cases here, it was the district that implemented a primarily control oriented, external goal oriented structural relationship, while in the other case, it was the high school that implemented a primarily commitment oriented, internal goal oriented structural relationship. Yet, even in the second case, the district during the time of this study began to exercise more control for external purposes.

Research Problem

We know from largely independent bodies of research that teacher evaluation systems and PLC models widely diverge in theory and practice. Whereas, teacher evaluation systems produce individual summative ratings for employment decisions, perhaps ideally along with formative feedback (e.g., Papay, 2012), PLCs aim to foster individual and collective capacity through collaborative professional learning, perhaps ideally also fostering professional accountability (e.g., Kruse, et al., 1995). Evaluation relies on personal accountability and bureaucratic control for its intended effects, PLCs on collective learning and professional commitment. They seem to stand at opposite poles along a number of education policy dimensions, for example, from governance centralization versus decentralization to bureaucratic versus professional accountability, to individual versus collective responsibility, learning or capacity (see Adams and Kirst, 1999; Bryk, Sebring, Allensworth, Luppescu, and Easton, 2010; Elmore, Ableman, Even, Kenyon, and Marshall, 2004; Ingersoll, 2003; Ingersoll and Collins,

2017). PLC models, in short, exemplify the paradigmatic commitment oriented design, performance evaluation systems the paradigmatic control oriented design (Rowan, 1990; see Darling-Hammond, et al., 1983; Ingersoll, 2003).

To adopt both in the same district or school absent attention to how the policies converge on teachers in implementation seems like a recipe for incoherence. Two decades ago, Knapp, et al. (1998, p. 398) identified the real-world problem that educators in countless districts and schools have come to share:

> The fact of multiple reforms begs questions about their interactions with one another and about *their joint influence on the working circumstances of professional educators and ultimately on the quality of learning opportunities available to children*. There is no guarantee that, in combination, the reform initiatives reinforce each other. It is just as possible that *multiple reforms, separately conceived, might get in each other's way or simply overload the system* … so that little, if anything, is accomplished.

(Emphasis supplied; see also, Srinivasan and Archer, 2018). Since all kinds of districts/schools, facing all kinds of pressures and challenges, are subject to multiple reforms, the potential for deleterious policy interactions, joint influences, and uncertain effects in implementation will be realized in otherwise disparate district, school and educator contexts (see, e.g., Louis and Robinson, 2012; Marsh, et al., 2017). Further, it is not just district and school administration that must confront the resulting coherence challenges (Honig and Hatch, 2004). Rather, perforce the organization of schooling, "[i]t remains to the frontline professional to sort this all out …" (Knapp, et al., 1998, p. 409). Unless mitigated by effective, ongoing "crafting" in which teachers have a large stake and to which teachers could contribute, incoherence bodes ill for teacher quality improvements and for reform—that is, school and student—success.

PLC models and teacher evaluation systems are widely implemented reforms of general applicability. How they were "separately conceived" and diffused historically can be another

problematic factor in implementation. While promising improvement across school contexts, each reform has taken an independent route through development, adoption, and implementation. Many districts began adopting a PLC model in response to school level accountability mandates like those culminating in 2001 with the No Child Left Behind Act (NCLB) during the second Bush administration. The successor Obama administration used its NCLB waiver authority and Race to the Top program (RTTT) to induce many states to enact teacher evaluation systems (Lavigne, 2014; Pogodzinski, Umpstead and Witt, 2015; Reinhorn, Johnson, and Simon, 2017). In many districts, then, an unprecedented external system of teacher evaluation was layered onto an existing internal system of PLCs (Jones, Bettini, and Brownwell, 2016; Woodward and Mazur, 2015). A similar layering, just in reverse time sequence, occurred in those districts that adopted a PLC model after the advent of state-mandated evaluation.

The consequences of policy convergence in either sequence have been similar. For as the two policies converge on schools, their disparate demands interact to affect teacher responses (Knapp, et al., 1998; Rigby, et al., 2016). After all, teachers' responses to one must come in the context of their responses to the other, particularly because both involve some of the same educational elements, activities and artifacts. Thus, a teacher's response to a personalized accountability demand can affect the same teacher's response to a demand for collective improvement work, and vice versa (Harris and Herrington, 2015; see Jones, et al., 2016; Woodland and Mazur, 2015). Teacher responses together redound on the policy structures to affect their ongoing implementation (Burch, 2007; Coburn, 2016). Following initial implementation, moreover, teachers (and administrators) re-enact both teacher evaluations and PLCs every school year, even as both interact and evolve in the process. The dynamic of

iterative, routinized implementation over time leads to the institutionalization of both policies (Coburn, 2016; see Burch, 2007).

The challenge of attaining, let along maintaining, policy coherence does not end with institutionalization of course. To the contrary, it persists indefinitely, not only because the institutionalized structures themselves continue to change, not only because there are similarities or overlaps in some elements available for use in response, but fundamentally because teachers each school year must engage anew with the related structures of PLC and evaluation in order to respond to their particular demands. Those demands specify goals, activities, tasks and products that may or may not aid efforts to achieve coherence in teacher responses and effects. Moreover, PLC and evaluation demands, whether separate or interacting, form structures in the work environment that teachers must navigate, while yet pursuing the primary occupational objective: successful teaching and learning. How teachers understand and act on the demand structures and their responsive agency within them, as well as how teachers respond and why, present empirical issues this dissertation addresses.

Research Questions

Each school year, districts and schools implement evaluation and PLC through established structures. Although the structures are not entirely static—they are related, and modifications are introduced—teachers are subject to the concurrent demands of both at various points over the school year. This research aims to understand and explain how teachers' agentic responses to these dual demands might affect each other, and thereby affect how teacher deal with both evaluation and PLC in practice. Relatedly, this research examines how the characteristics of the structures of evaluation and PLC can upon interaction affect the nature of

and scope for teachers to perceive opportunities and to take action to advantageously adapt their responses for their own ends.

Specifically, this study addresses the following research questions:

1. How, if at all, do teachers adapt their agentic responses to performance evaluation and professional learning community demands as they respond to the structures of both?

2. How, if at all, does the relationship between structures of performance evaluation and professional learning community shape teachers' adaptive responses?

These research questions underline the significance of two subjects of this study: teachers' "agentic responses" to the demands of PLC and evaluation, and the "structures" of PLC and evaluation. Chapter 3 will discuss structure-agency theory in more detail, but for definitional purposes now, "agency" concerns a person's capacity to act (Donaldson and Woulfin, 2018); as Colburn (2016, p. 466, quoting Scott, 2008, p. 78), explains: "Agency can be understood as 'an actor's ability to have some effect on the social world—altering rules, relational ties, or distribution of resources.'" Coburn (2016, p. 466) also adopts Scott's (2008, p. 78) definition of social structure as "'the patterning of social activities and relations through time and across space … incorporating rules, relations and resources reproduced over time.'" Structures support, direct and guide organizational activities, specifying how organizations function and how actors perform their roles (Donaldson and Woulfin, 2018; Rigby, et al., 2016; Sewell, 1992).

Taking a structure-agency theoretical approach to frame the research questions and guide the analysis of the relationship of *two* teacher agentic responses to *two* distinct demands within *two* related structures may be novel with this dissertation.

Dissertation Contribution

This research is an empirical investigation of how teachers adapt their otherwise independently elicited responses in order to navigate the demands of these two disparate in principle, if interacting in practice, teacher quality policies. Through the identification and explanation of both elements of this teacher two-step—two agentic responses to two structural demands—this dissertation makes the following contributions.

First, this dissertation contributes to a body of research showing the utility and versatility of a structure-agency conceptual framework (see Coburn, 2016; Rigby, et al., 2016; e.g., Donaldson and Woulfin, 2018). It is practically a truism that actors charged with policy implementation will in that very process change the policy as designed. Here, a structure-agency lens has been turned on an ongoing coherence challenge: two very different teacher quality improvement structures are re-enacted together with the same policy targets year upon year. The convergence of similar ends, differing means, and shared periods of reproduction results in a complex of inter-relationships among two agentic responses, which can influence each other, and two structures, which can interact with each other. The resulting analysis and findings provide insights into and suggest explanations for the patterns and variations among adaptive responses as shaped by related structures. This dissertation thus represents a fruitful, knowledge building extension of the structure-agency approach to a new, if common, education policy context.

Second, this dissertation suggests that a broader, more elaborated conception of the participants who engage in and mechanisms constituting "crafting coherence" may better guide local policymakers and practitioners in their efforts to achieve coherence among current and future teacher quality policies. The results here confirm Honig and Hatch (2008) in two main respects: policy incoherence poses an ongoing challenge to school improvement efforts; and

conversely, policy coherence must be constantly made and re-made through a dynamic process of adjusting responses to internal and external demands. Moreover, each administration in the cases here intentionally undertook to implement their evaluation and PLC structures in relation to each other; and although their measures differed, they were both nonetheless encompassed within Honig and Hatch's (2008) conceptualization (see also DiPaola and Tschannen-Moran, 2005). At the same time, however, by showing that teachers in PLCs can and do participate in their own form of coherence crafting, this dissertation extends to the bottom or "street level" our understanding of who crafts coherence and how. Teacher crafting in this study involves dynamically adapting responses through the application of distinct strategies developed within the related structures of evaluation and PLC. And while the implementation of related structural, demands influence teacher responses; attaining coherence all the way down still depends upon the exercise of teacher agency (see Russell and Bray, 2013; Stosich, 2018).

Finally, even as this dissertation will begin to fill a gap in the literature, it has implications for researchers, policymakers, and practitioners. As previously suggested, despite researcher interest in discrete teacher quality reforms, especially when each is new, newly popular, or initially implemented, little literature explicitly investigates the responses of teachers, alone or in formalized groups, to the sustained convergence of two performance improvement and accountability demands quite like evaluation and PLC. Nor does much extant research concern either why teachers respond to certain aspects of policy convergence, or how teachers participate in coherence crafting in the process of response. This dissertation starts to bridge the largely separate research silos by adding to our knowledge of the interaction of policy demands and the relationship of policy target responses in the routinized implementation of policy prescriptions. Further work building on this research might sufficiently deepen and broaden our

understanding of teacher responses to related policy demands to inform multiple policy design implementation, institutionalization and evaluation.

This research helps explain how teacher agentic response to each "new new thing" in teacher quality reform, regardless of source or intention, will from the bottom up impact the coherence crafting efforts among the existing constellation of policies at the school and district levels. One implication for policymakers and practitioners is therefore this: even apart from a policy's theory of action and fidelity of execution, one teacher quality policy can alter the functioning and results of another. Recursively driven by the responses of the policy targets themselves, as shaped by local design and goal orientations, such downstream effects should be anticipated and accounted for in the design and implementation of convergent reform proposals. This and subsequent research therefore have the potential to inform the consideration of any new or modified policy that would likely interact with existing systems and influence teacher responses to multiple demands.

From the perspective of the policy targets that are this study's primary subjects, this research illuminates why teachers neither need nor should navigate multiple teacher quality structures via rigidly routinized, discrete responses of mainly resistance or acceptance. Instead, teachers can and should with intention leverage their individual and collective agency to expand the scope and range of effective interacting responses. Especially if pursued in win-win concert with administration, teachers might then promote the utility of PLC work and evaluation, while also affirming teacher professionalism and professional accountability in the service of their own professional growth.

Dissertation Outline

The chapters that follow review the relevant literature, describe the evaluation/PLC policy context and structure-agency conceptual framework, set forth the qualitative research design and methodology, explain the findings as to the first and second research questions, and discuss those findings and their implications.

The first section of *Chapter 2* describes the research on the nature and characteristics of PLCs, a commitment oriented teacher quality policy, including how PLC demands and teacher responses influence PLC functioning and results. The second section takes the same approach to elucidating the research on teacher performance evaluation as a control oriented teacher quality policy. Finally, a third section reviews the limited research bearing directly on the interaction effects of one policy on the other.

Chapter 3 first provides the policy context for this study. It describes the enactment and terms of Michigan's teacher evaluation system, as well as the nature and adoption of PLC models in many Michigan districts. The second section of Chapter 3 elaborates on why the PLC/evaluation policy context poses a challenge of coherence to educators subject to both each school year. Given this challenge, the second section proceeds to present the components and relationships of the structure-agency approach that has been applied to address this dissertation's research questions.

Chapter 4 in its first section delineates the multiple case study design that this research employs, together with its rationale. The research design discussion is followed in succeeding sections by explanations of this study's sampling strategy and participants, data collection and analysis, and procedures undertaken to establish validity.

Chapters 5 and 6 present the evidence and findings that answer, respectively, the first and second research questions. Chapter 5 focuses on the teacher agency component of the structure-agency approach. Following a description of the overall evaluation/PLC relationship as one characterized by "punctuated compartmentalization" of demands and responses, the chapter goes on to demonstrate when, how and why teachers adapt their responses to the demands of evaluation based on their PLC work directed toward achieving particular PLC objectives. Chapter 5 therefore identifies and describes four adaptive response strategies teachers may use (and did use) to their advantage in dealing with related evaluation and PLC demands. It also compares the strategies along two intersecting dimensions.

Chapter 6 concerns the structural side of the conceptual approach. The two districts/ schools implemented related PLC and evaluation systems with differing design and goal orientations. One used a control orientation as a means to an external goal, while the other's commitment orientation aimed at an internal goal. This chapter shows how these alternative paired orientations in implementation influenced teachers' use of various adaptive response strategies.

Finally, *Chapter 7* concerns the broader significance and contributions of this research. This study, first, enhances our knowledge and understanding of the structure-agency approach; and next, how coherence may be crafted and by whom. This includes types of teacher agentic responses driving implementation within related or interacting structures that vary in their orientations, and the implications of teachers' participation in such crafting. The chapter next explores implications for policymakers, practitioners, and researchers. It wraps up with a discussion of the limitations of this research and a brief section in conclusion.

CHAPTER 2
LITERATURE REVIEW

The two teacher quality reforms this study concerns originated at different policymaking levels, at different times, in different places, and in different policymaking and implementation contexts. Broadly speaking, PLC and evaluation policies are opposed in design elements—the former representing a strongly commitment oriented model, the latter a strongly control oriented one (Rowan, 1990; see Darling-Hammond, et al., 1983; Ingersoll, 2003). Not surprisingly, then, the development and spread of PLC models and teacher evaluation systems spawned substantial but largely independent bodies of research. Nonetheless, we know that many districts and schools across the United States have established and annually enact both.

This chapter has four sections. The first two concern those separate bodies of PLC and evaluation research. These sections undertake to describe and synthesize that portion of each body of research that might make coherence challenging when both reforms interact or converge on schools and teachers. The discussion for each therefore focuses on policy theories of action and goals, together with, in implementation, each reform's demands and influences, and teacher responses and effects. The third section reviews the far more limited research, theoretical and empirical, that involves substantial interactions between evaluation systems and PLCs, together with the resulting variation in school and teacher responses. This chapter's concluding section summarizes the overall nature of the research and identifies a gap this research aims to begin filling.

PLC: A Commitment Oriented Teacher Quality Reform

Although the development of PLCs seems many-sourced, a body of professional community research becomes discernable beginning in the 1980s (Stoll, Bolam, McMahon, Wallace, and Thomas, 2006). A movement for PLCs (under a variety of monikers) gained

momentum in the 1990s alongside school-level accountability and state standards reforms. It grew and spread in the 2000s with the passage and implementation of NCLB. From the first to the second decade of the new millennium, many districts and schools around the nation adopted formalized teacher professional community or collaborative team models (Blankenship and Ruona, 2007; Brouwer, Brekelmans, Nieuwenhuis, and Simons, 2012; Charner-Laird, Ng, Johnson, Kraft, Papay, and Reinhorn, 2016; Feger and Arruda, 2008; Levine, 2010). Today, despite a plethora of names and configurations (Levine, 2010; Murphy and Torre, 2014; Stoll, et al., 2006), PLC models tend toward a common theory of action and similar design features inherited from earlier practitioner experience and academic research (Brouwer, et al., 2012; Hord, 1997; Kruse, Louis and Bryk, 1995; Levine, 2010; Murphy and Torre, 2014; Stoll, et al., 2006).

PLCs are clearly a commitment model of school improvement (Rowan, 1990). "At the heart of the [PLC] concept … is the notion of community. The focus is not just on individual teachers' professional learning but on professional learning within a community context—a community of learners, and the notion of collective learning" (Stoll, et al., 2006, p. 225). The description Hord (1997, p. 6) gives likewise emphasizes collective learning, but it points to action and an objective. She says that, in a professional community of learners,

> … the teachers in a school and its administrators continuously seek and share learning, and act on their learning. The goal of their actions is to enhance their effectiveness as professionals for the students' benefit; …

(See also Stoll, et al. 2006, p. 223; Murphy and Torre, 2014, p. 101). So constituted, PLCs are designed to leverage teacher knowledge, skill and experience in a collaborative setting to foster individual and collective learning. Such professional learning will in turn build teacher and school capacity, improve teaching and learning, and satisfy internal and external accountability

demands. “Because of their inherent learning potential, communities of practice offer a pathway to improved teacher quality and educational change” (Printy, 2008, p. 190).

Research has specified the attributes of effective, well-functioning PLCs (e.g., literature reviews, Feger and Arruda, 2008; Hord, 1997; Murphy and Torre, 2014; Stoll, et al., 2006). The pioneering work of Kruse, Louis, and Bryk (1995) proposed a framework for analyzing school-based professional community according to five characteristics: shared norms and values, reflective dialogue, deprivatization of practice, collective focus on student learning, and collaboration (see also Louis, Kruse, and Bryk, 1995; Louis, Kruse, and Associates, 1995; Louis, Marks and Kruse, 1996; Louis and Marks, 1998). Stoll et al. (2006) later derived five mutually reinforcing characteristics from a review of the literature: shared values and vision, collective responsibility, reflective professional inquiry, collaboration, and group as well as individual learning. Based on her decade of research focusing creating successful PLCs in schools, Talbert (2010) described four core features based on how PLCs (1) establish expectations and norms for collaboration; (2) focus PLC work on improving all students’ achievement; (3) create knowledge resources for PLCs; and (4) engender mutual accountability for success. Finally, Murphy and Torre (2014) surveyed the literature for what they called the core elements of communities of professional practice. As the foundation of their generalized PLC model, Murphy and Torre (2014) argued for six defining elements: shared vision, collaboration, ownership, shared leadership, shared accountability, and trust. Crucially, Murphy and Torre (2014) emphasized, PLCs sharing these elements develop and practice collective responsibility and accountability.

Murphy and Torre (2014) in their literature review explicitly conclude that scholars’ findings of defining PLC attributes imply effective professional accountability as a hallmark of true PLCs (see also Talbert, 2010). The research does demonstrate that the attitudes, norms and

practices constituting professional accountability can emerge through teachers' engagement in PLCs (Johnson, Reinhorn, and Simon, 2018; Louis, Kruse, and Bryk 1995; Louis, et al., 1996; McLaughlin and Talbert, 2001; Murphy and Torre, 2014; Talbert and McLaughlin, 1994).

The literature on accountability in education policy has long recognized and described the characteristics of professional accountability in terms that closely correspond to the PLC characteristics discussed above. Thirty years ago, for example, Darling-Hammond (1989, p. 67) summarized three principles of professional accountability comprising (1) the authority to practice based on specialized knowledge applied to students' unique needs, (2) a pledge of first concern for students, and (3) the assumption of "collective responsibility for the definition, transmittal, and enforcement of professional standards of practice and ethics" (1989, p. 67; see Darling-Hammond, 1990; Adams and Kirst, 1999; Darling-Hammond and McLaughlin, 1999; Firestone and Shipps, 2005; Leithwood and Earl, 2000; Mayer, 2005; Mehta, 2011; O'Day, 2008; Sykes, 1999). Teacher engagement in professional community can lead to professional accountability when teachers share an accurate understanding of the most appropriate educational means and ends for their students. Further, teachers must hold each other to current standards of expert knowledge, skill and judgment to properly enact the agreed means to achieve the desired ends (Adams and Kirst, 1999; Darling-Hammond, 1989; Darling-Hammond and Ascher, 1991; Firestone and Shipps, 2005; Mehta, 2013; McDermott, 2011). Put somewhat differently, teachers' collective responsibility for learning is an essential element of professional culture with accountability (Lee and Smith, 1996; see O'Day, 2002, 2008). Møller (2008, p. 40 [author's emphasis]) seems to directly channel the PLC literature when he concisely sums up the point:

> … *[P]rofessional accountability* [is] where a person's commitment to a community of professionals makes him/her perceive a duty to

> adhere to the standards of the profession. … Professional accountability implies that teachers acquire and apply knowledge and skills needed for successful practice. In addition, it involves the norms of putting the needs of students at the centre of their work, collaborating and sharing of knowledge, and a commitment to the improvement of practice.

In addition to specifying the determinants of PLC performance, the literature has investigated factors that support (if present) or hinder (if absent) the cultivation and growth of PLCs (see Brouwer, et al., 2012). The analytical framework proposed by Kruse, et al. (1995), for example, included five structural conditions (e.g., time to meet, teacher empowerment and school autonomy) and five social/human resources (e.g., openness to improvement, trust and respect, supportive leadership) that support professional community at the school level. Similarly, Stoll, et al. (2006) from the author's own research added three contextual factors to those PLC characteristics identified in the literature (listed above): mutual trust, respect and support among staff members; inclusive membership extending beyond teachers and school leaders; and seeking sources of learning and ideas outside the school (Stoll, et al., 2006).

Research in line with Kruse, et al. (1995) and Stoll, et al. (2006) has likewise extended and refined the organizational conditions and supports that promote effective PLC functioning and sustainability in a variety of forms. Such factors span a range from structural, such as formal opportunities, quality resources, ongoing professional learning and participatory leadership structures to support authentic collaboration, to more cultural components, such as shared goals, an inquiry focus, and broad-based trust, to committed and effective district and school leadership (DuFour and Fullan, 2013; Louis, Dretzke, and Wahlstrom, 2010; McLaughlin and Talbert, 2001; Murphy and Torre, 2014; Supovitz, 2002; Szczesiul and Huizenga, 2015; Tschannen-Moran, 2009; Wood, 2007). Additional research has focused on those resources in the broader context that can advance PLC work. Researchers have examined the nature and extent of

relational trust (Bryk and Schneider, 2001; Miesner, Blair, Packard, Velazquez, Macgregor, and Grodsky, 2019) and collective trust (Forsythe, Adams, and Hoy, 2011), as well as features of the school environment such as the presence and substance of collective professional norms (Murphy and Torre, 2014; Kraft and Papay, 2014), an organizational learning culture (Louis and Lee, 2016) or culture of learning (Tichnor-Wagner, Harrison and Cohen-Vogel, 2016), and professional capital (Fullan and Hargreaves, 2015).

It is plain from the multiplicity of multifarious sets of attributes, conditions and supports that successful PLCs impose high demands on educators and schools (DuFour and Fullan, 2013; Hargreaves, 2007; Murphy and Torre, 2014). This fact, together with challenges in theoretical conceptualization and empirical operationalization for empirical research, may help explain the suggestive but limited recent evidence linking PLCs to improvements in teacher practice and student achievement (Akiba and Liang, 2014; Lomoa, Hofman, and Bosker, 2010; Ronfeldt, Farmer, McQueen, and Grission, 2015; Vescio, Ross, and Adams, 2008). PLCs' high practitioner demands might likewise help explain what researchers (let alone practitioners) have long recognized: schools vary widely in how well actual PLCs develop and function to fulfill their promise (Charner-Laird, Ng, Johnson, Kraft, Papay, and Reinhorn, 2016; Hargreaves, 2007; Hargreaves and Dawe, 1990; Johnson, Reinhorn, and Simon, 2018; Louis, Marks and Kruse, 1996; McLaughlin and Talbert, 1993; Ronfeldt, et al., 2015; Supovitz, 2002; Woodland and Mazur, 2015). Much of the literature seeking to understand the variation in PLC performance focuses on two factors relevant to the instant study. One, the role of school leadership in establishing, operating and supervising PLCs, concerns the demand side of implementation, whereas the other, teachers' motivation and capacity to engage in PLCs, concerns the response side.

School level leadership can act to improve an array of teacher and organizational capacities (Bryk, et al., 2010), including shared or distributed leadership (Louis, et al., 2010), relational trust (Bryk and Schneider, 2002; Tschannen-Moran, 2009), shared responsibility within internal accountability (Elmore, 2005), and a focus on school goals with active support of instructional improvement (Louis, et al., 2010; Supovitz, et al., 2010). In addition to its role in developing PLC supports like those previously identified (see, generally, Murphy and Torre, 2014), school leadership, especially principals, can more directly impact PLC performance (Buttram and Farley-Ripple, 2016; Charner-Laird, Ng, Johnson, Kraft, Papay, and Reinhorn, 2016; DuFour and Eaker, 1998; Feger and Arruda, 2008; Johnson, et al., 2018; Murphy and Torre, 2014; Printy, 2008; Wood, 2007; Young, 2006).

Some of this literature has tried to disentangle somewhat more specifically those behaviors by which formal school leaders influence teachers in PLCs. For example, Buttram and Farley-Ripple (2016) in their study of principals implementing a new state-mandate for "PLC time" in schools found that the participants' understanding of the mandate and strategies for its implementation strongly influenced what teachers did in PLCs and how well they did it. Charner-Laird, et al. (2016) examined the performance of teacher teams in six high-poverty schools from the teacher-participants point of view. The researchers found substantial variation among schools in how well PLCs functioned, with an overall determination that the PLCs in three schools were a "good fit" for teacher and school learning and three a "bad fit." The authors detailed their findings for each school, but the highest level factors they focused on consisted of how successful the principals were in "defining the purpose of teams, setting the environments in which they worked, and supporting their work along the way," concluding that "it appears that these formal school leaders [principals] will continue to be crucial partners in determining how

and how well teams work in schools" (Charner-Laird, 2016, p. 33) Finally, Printy (2008) studied how secondary principals and department chairs exerted influence on the composition and quality of teacher participation in PLCs, for good or ill. She suggested that successful leaders engaged in several activities, from creating the conditions for effective interactions for learning to using guided activities to structure meetings to modeling sensemaking as teachers learn to scaffold their knowledge. The most critical influence seemed to stem from the expectations school leaders established for teacher participation (see also Datnow, 2011; Johnson, et al., 2018; Young, 2006).

How school leaders design, implement, and oversee the overarching approach to and structures of PLCs can strongly shape teacher response to PLC demands and so the outcomes of PLC work (Wood, 2007). Thirty years ago, Hargreaves and Dawe (1990, p. 230) contrasted

> …how far some current efforts toward developing norms of collegiality among teachers create critical and *collaborative teacher cultures* which develop curriculum and pedagogical reform from within the profession, on the one hand, or forms of *contrived collegiality*, which are administratively designed to smooth the path of externally imposed innovation, on the other.

(Emphasis in original; see also Datnow, 2011; Hargreaves, 1994). Writing twenty years later, Talbert (2010, p. 555) in a "sobering assessment of the challenges facing the PLC movement" updated and elaborated the problem for formalized systems of PLCs in her comparison of two strategies for their implementation.

Talbert (2010) designates one strategy "bureaucratic," the other "professional," arguing that "[t]he strategies a school district uses to create PLCs make a great difference in outcomes" (Talbert, 2010, p. 560). The bureaucratic approach rests on principles of top-down control, where schools implement PLCs through such measures as requiring teachers to collaborate, establishing leadership roles, and making PLCs accountable up the command chain. The professional

approach is much more commitment oriented, relying on processes like developing a common vision, increasing capacities for leaders to support change and for educators to address gaps in student achievement, and creating a set of inter-related PLC resources. Critically, according to Talbert, teacher responses to bureaucratic demands usually involve compliance, resistance and anxiety, while professional demands tend toward enthusiasm, cooperation and wait-and-see collaboration. Still, the two "typically exist and compete in school system [sic] where leaders embrace different views of how to create PLCs" (Talbert, 2010, p. 561; see Charner-Laird, 2016). Since PLCs can only sustainably succeed when they work change in a school's professional culture, and not when adopted as another school improvement "program," Hargreaves (2007), Datnow (2011), and DuFour and Fullan (2013) effectively join Talbert (2010) in urging the professional strategy, even as they use different verbal formulations of it.

On the other side of PLC implementation, teachers within intra-school and particular PLC contexts may lack sufficient motivation or capacity to respond effectively. First of all, there are likely persistent norms that may be inimical to those upon which authentic PLCs depend. Strong norms of teacher autonomy and equality, managerial imperative, and civility may at best hinder the development of professional learning and development of professional accountability (Murphy and Torre, 2014; see also DuFour, 2004; Little, 1982, 1990; Lortie, 2002 [1975]). These norms, moreover, can combine with poor teacher discourse into So-called "collaboration light" (DuFour, et al., 2004; see Wood, 2007)—unfocused, superficial, negative—or worse (see, e.g., Datnow, 2011)—can lead to unproductive at best and toxic at worse PLC dysfunction that ends up reinforcing teachers' pre-existing practices (see Horn and Little, 2010; Talbert, 2010; Woodland and Mazur, 2015).

Still, to the extent a critical mass of teachers in PLCs is able to develop opposing norms and engage in productive discourse such as those discussed at the beginning of this section (e.g., Kruse, et al., 1995), then deprivatized practice and authentic collaboration may emerge, perhaps bolstered by effective systems of support (Louis and Lee, 2016; Johnson, et al., 2018). In the movement from one state to another, whether achieved or not, seemingly unavoidable conflict may or may not be successfully navigated, with effects differing among teachers and schools (Achinstein, 2002; Johnson, 2003; see Datnow, 2011).

The first challenge to the productive functioning of PLCs points to a second: substantial variation in teacher capacities to engage quality collaboration (Charner-Laird, 2016; Rongeldt, et al., 2015; Supovitz, 2002; Wood, 2007). One research approach to discourse capacity question has been to identify those factors positively influencing the work of PLCs overall. For example, several studies showed that, beyond PLC supports like sufficient meeting time and active administrative support, the use of trained facilitators contributes to PLC success (Andrews-Larson, Wilson, and Larbi-Cherif, 2017; Bolhuis, Schildkamp, and Voogt, 2016; Reinhorn, et al., 2018). Other researchers have looked to the purpose of PLC collaboration. Popp and Goldman (2016) compared two sets of grade-level PLC meetings—one teacher collaboration focused on assessment systems, the other on instructional activities—to determine in which there were more participant moves associated with knowledge building discourse. The authors found that the former, assessment systems discourse was superior, and more generally suggested that "meeting foci play a role in supporting (or not supporting) knowledge building in teacher PLCs" (Popp and Goldman, 2016, p. 356). Similarly, in their literature review, Vescio, et al., (2008, p. 89) concluded that a key commonality among the cases of successful PLCs was "collaboration with a clear and persistent focus on data about student learning" (see also Ronfeldt, et al., 2015; Young,

2006). The literature also includes support for the significance of collaborative foci such as to change teachers' instructional practices (Supovitz, 2002) and to serve a worthwhile purpose in support of the school's mission (Reinhorn, et al., 2018).

Another approach of scholarly attention concerned with teacher capacity for professional learning through collaboration investigates variations in how teachers actually interact in PLCs. Typically based on the fine-grained analysis of "teacher talk," this research works from the bottom up to understand how teacher discourse produces learning opportunities and resources that support improvements in teacher practice (Horn, Garner, Kane, Brasel, 2017; Horn and Little, 2010; Little, 2002a, 2003; Richmond and Manokore, 2010). Horn and Little (2010, p. 183) studied how the conversational routines in two teacher work groups in the same school "enhanced or limited opportunities for the in-depth examination of problems of practice and hence shaped opportunities for teacher learning." Although the two seemed alike in many ways, sharing a commitment to improving teacher practice and a school context, "these groups were found to differ in the opportunities for learning they constructed through micro-level discourse routines they employed … and the meso-level participation routines they used…" (Horn and Little, 2010, p. 211). Along with each group's orientation and contextual resources, it was the qualitative nature of the discourse that produced the potential for learning. In more recent work, Horn and colleagues studied mathematics workgroups across sixteen middle schools that were selected as exemplars of teacher collaboration (Horn, et al., 2017; see also Horn, Kane, and Garner, 2018). Seeking "to uncover how different kinds of collegial conversations shape teachers' professional learning opportunities" (Horn, et al., 2017, 41), the researchers developed and tested a taxonomy of teacher team meetings based on the extent and nature of the teacher learning each kind constituted. Ranging from least effective to most effective as generative of

opportunities for authentic learning, the six kinds were (1) conflicting goals, (2) pacing, (3) logistics, (4) tips and tricks, (5) collective interpretation separate from future work, and (6) collective interpretation linked to future work. This research shows how important authentic, quality collaboration, consisting largely of teacher discourse, is to both the promise and challenge of establishing and sustaining effective PLCs.

Little (2003) in a sense captures much about the potential for interactions in PLC to improve learning and practice. After summarizing her examination of "the intellectual, social, and material resources of professional community, particularly as those are constituted through representations of classroom practice in out-of-classroom [PLC] interactions" (Little, 2003, p. 938; note omitted), she concludes that the resulting findings

> …complicate some of the existing typologies of teacher community by showing how *ongoing interactions both open up and close off opportunities for teacher learning* and consideration of practice—in the same groups and sometimes the same moments. Even within these groups that would be reasonably considered collaborative, innovative, and committed to improving practice, teacher *learning seems both enabled and constrained* by the ways teachers go about their work. *The forces of tradition and lure of innovation seem simultaneously and complexly at play in teachers' everyday talk.*

(Little, 2003, p. 941; emphasis supplied; see Wood, 2007).

High-functioning PLCs are hard to actualize and sustain (e.g., DuFour and Fullan, 2013; Hargreaves, 2007; McLaughlin and Talbert, 2001). Nevertheless, PLCs, perhaps in modified or extended forms, still hold promise for advancing teacher quality in schools (Bryk, Gomez, Grunow, and LeMahieu, 2016; Quintero, 2017; Woodland and Mazur, 2015). Indeed, Bryk, et al. (2016) re-envisions teacher collaborative teams as Networked Improvement Communities (NICs) run according to six improvement principles (Bryk, et al., 2016). PLCs can at once build the individual and collective professional learning and capacity (Newmann, et al., 1997;

Charner-Laird, 2016) that promotes professional capital (Fullen and Hargreaves, 2015) and internal accountability (Elmore, et al., 2004) for student learning and school improvement (McLaughlin and Talbert, 2001; Ronfeldt, et al., 2015; Louis and Marks, 1998).

Teacher Evaluation: A Control Oriented Teacher Quality Reform

Following on years of policy ferment about the inadequacies of local district teacher evaluation processes (e.g., Donaldson, 2009; Weisberg, Sexton, Mulhern, and Keeling, 2009), the RTTT and NCLB waiver policies of the Obama Administration strongly prompted states to adopt high-stakes, teacher evaluation as a major teacher quality reform (Donaldson, 2009; Kraft and Gilmour, 2016; Lavigne, 2014; McGuinn, 2012; see National Council on Teacher Quality, 2017; Reinhold and Johnson, 2015). Most state-mandated evaluation systems as enacted seek the achievement of two goals: a *summative* purpose lying in teacher accountability and a *formative* purpose lying in teacher development (Darling-Hammond, 2013; Papay, 2012; Kraft and Gilmour, 2015; Reinhold and Johnson, 2015; see, e.g., Dee and Wyckoff, 2015 [District of Columbia]; Pogodzinski, Umpstead,and Witt, 2015 [Michgian]; Reinhorn, Johnson, and Simon, 2017 [Massachusetts]).

As to the accountability or summative purpose, "the core of teacher evaluation is what it is, a system of performance management that is scaffolded on hierarchical authority and control over labor [references omitted]" (Murphy, Hallinger, and Heck, 2013). This purpose relies on a bureaucratic accountability design for its theory of action (Adams and Kirst, 1999; Darling-Hammond, Wise, and Pease, 1983; Maslow and Kelley, 2012). Taking a classic control approach (Rowan, 1990), the system primarily aims to hold teachers individually accountable for their teaching quality and their students' learning by using multiple measures of effectiveness. Its key components are evaluators' ratings of standards-based classroom observations and the

determination of student achievement growth from standardized assessment scores (Goe, Bell, and Little, 2008; Woodland and Mazur, 2015; Papay, 2012; Whitehurst, Chingos, and Lindquist, 2014). Assuming that the components together produce a valid and reliable effectiveness rating for each teacher, teachers can be ranked for purposes of high-stakes personnel decisions on teacher retention (lay-off, non-renewal, discharge), promotion or additional compensation, and the award of tenure or other job security benefits (Hallinger, Heck, and Murphy, 2014; Papay, 2012).

The potentially punitive consequences (or, sometimes, rewards) attached to evaluation effectiveness ratings are intended to provide powerful incentives for teachers to achieve high effectiveness or exit the profession (Firestone, 2014; Harris and Harrington, 2015; Papay, 2012;). The exiting of ineffective teachers, whether voluntarily or not (Dee and Wycoff, 2015; Loeb, Miller, and Wyckoff, 2015), should over time increase the overall level of teacher quality in a school or district (Papay, 2012; Hallinger, et al., 2014; Lavigne, 2014). In short, as Harris and Herrington (2015, p. 72) put it:

> The main underlying theory of these policies is that teacher accountability will motivate teachers to work harder and smarter and help attract and retain only those who are successful.

The professional development or formative purpose of teacher evaluation principally relies on the system's provision for individualized feedback on classroom observations. The student growth component comes too late in the process and in any event does not indicate *how* the teacher can or should improve their performance (Goe, et al., 2009; Whitehurst, et al., 2014). Kraft and Gilmour (2016, p. 715) locate the theory of action for formative evaluation in "cycles of observation, reflection, dialogue and feedback, and goal setting [that] can provide teachers

with new ideas as well as frequent and relevant feedback to support their professional growth" (see also Darling-Hammond, 2013; Papay, 2012; Reinhorn, et al., 2017).

It is not difficult to notice that the "[g]oals of development and accountability in teacher evaluation often stand in tension, even though policies governing evaluation typically call for achieving both" (Reinhorn and Johnson, 2015, p. 3; see also Darling-Hammond, 2013; Papay 2012; Ray Taylor and Associates, 2017). Hence, even before practitioners and researchers could address the usual challenges attending implementation of such a complex, resource-heavy, game-changing policy, its very design was highly problematic. Further, despite the fact that formative use of evaluation held greater promise for improvements in teaching and learning (Papay, 2012; Murphy, et al., 2013), Papay (2012, p. 124) pointed out that "[f]or the most part, policy debates regarding evaluation systems have revolved around the first of these purposes, focusing on evaluations as measurement instruments to assess teacher performance and hold teachers accountable" (see also Darling-Hammong, 2013; Woodland and Mazur, 2015). This fact has shaped the implementation research from both local policy demand and response perspectives.

The implementation of teacher evaluation systems has proved extremely challenging, due not merely to divergent goals, but also due to the high demands placed on and the responses required of local districts and schools (e.g., Dee and Wyckoff, 2015; Marsh, Bush-Mecenas, Strunk, Lincove, and Huguet, 2017; Reinhorn and Johnson, 2015). Effective implementation with fidelity meant that district and school personnel had or could acquire necessary time, capabilities, and other personal and organizational resources (see, e.g., Maslow and Kelley, 2012; Woodland and Mazur, 2015). As Kraft and Gilmour (2016, p. 715) succinctly stated with reference to one evaluation goal: "A key assumption of [the formative evaluation] theory of action is that teachers are both willing and able to improve their practice by actively engaging in

the evaluation process." The formative process further assumes that administrator/evaluators are willing and able to formulate, provide, and follow-up on prompt, actionable feedback to teacher/ evaluatees (Goe, et al., 2009; Kraft and Gilmour, 2016; Maslow and Kelley, 2012; Woodland and Mazur, 2015). Summative metrics involve largely isolated, perhaps unrepresentative, snap shots of teacher performance. They not only depend on a valid and reliable observation tool and student achievement data fairly linked to individual teachers, but also, again, on the competence of administrators to understand sufficiently and apply fairly the observation tool, student achievement measures, and supporting technology. Teachers must be able to accurately comprehend their evaluators' demands in order to effectively respond to them given their individualized contexts (Goe, et al., 2009; Hallinger, et al., 2014; Maslow and Kelley, 2012; Murphy, et al., 2013; Papay, 2012; Woodland and Mazur, 2015).

Given the overall aims and principal components of evaluation demands/responses in schools, at least for purposes of this dissertation, the implementation literature on the evaluation demand side had examined the behaviors of school leaders, especially principals, in their immediate contexts. Principals can for the entire teaching staffs frame and prioritize evaluation, communicating how they approach and value it (Donaldson and Woulfin, 2017; Reinhorn et al., 2017). They directly engage in the evaluation process with "their" particular teachers and within varying organizational constraints and affordances of "their" school. Principals plainly exercise meaningful, even substantial, agency when they enact evaluation (Donaldson and Woulfin, 2017; Marsh, et al., 2017; Reinhorn and Johnson, 2015), not the least in how they address the challenge of implementing the development and accountability goals (e.g., Reinhorn, et al., 2017; Tuytens and Devos, 2010). The interplay of principal agency and contextual factors can largely determine the extent to which evaluations are conducted with fidelity and thereby might attain policy goals

(Donaldson and Mavrogordato, 2018; Donaldson and Woulfin, 2018; Kraft and Gilmour, 2016, 2017; Reinhold and Johnson, 2015; Reinhold, et al., 2017). As always with the implementation of an education policy, especially when externally imposed, leadership matters (see Honig and Hatch, 2004; see also, e.g., Marsh, et al., 2017; Tuytens and Devos, 2010).

Accordingly, to explain empirically the principal's role in evaluation implementation, some researchers examine those cognitive processes that principals employ to know and decide how to implement state-mandated evaluation systems. Reinhorn, et al. (2017) studied six high-performing, high poverty schools, where they found that the "principals' views, priorities and approach to evaluation" had the greatest influence on their teachers' understanding that the purpose and effect of evaluation was desirable individual improvement (Reinhorn, et al., 2017, p. 401; see also Tuytens and Devos, 2010). Looking to sensemaking theory, Reinhorn, et al. (2017, p. 401) summarized the sources of the principals' implementation behaviors by stating that they:

> drew upon their knowledge and skills about good teaching and a commitment to use strategies that would support teachers' development (individual cognition). They also had a clear understanding of what their school's particular policy context encouraged and allowed, and the capitalized on the opportunities it provided (situated cognition). Furthermore, they recognized the state's policy stimuli, which highlighted the importance of teachers' continuous development, while not ignoring the importance of dismissing weak teachers. This focus, conveyed by the state's [evaluation documentation] and additional training and supports, aligned with these principals' professional priorities.

Studies somewhat more narrowly focused on particular participant groups or evaluative contexts have identified various considerations that inform principals' personal efforts to implement evaluation in their particular contexts. For example, Rigby's (2014) research interest involved the messages six first year principals received about the enactment of the same evaluation system components. The participating principals were chosen in part because they had attended a distinct principal preparation programs. Taking a sensemaking approach akin to

Reinhold, et al. (2017), Rigby found that, with respect to their roles as instructional leaders conducting evaluations, the principals were most influenced by the messages they connected to that originated from their informal collegial networks of other participants in their respective programs. This influence was key even as compared to the school context and previous experience. Although the primary target of Reinhold and Johnson (2015) concerned variation in how principals used evaluation for their own school-level purposes (discussed below), they also considered the potential impact of participant attributes, such as experience, professional learning, understanding and attitude with respect to principals' enactment of evaluation.

Donaldson and Mavrogordato (2018) studied 17 principals and assistant principals in ten schools to learn how they evaluated low-performing teachers to move toward improvement for retention or dismissal from employment. The researchers found that the principals' personal "framing of teacher performance and teacher ratings, and their efforts to improve instruction reveal the cognitive, relational and organizational aspects of working with low-performing teachers" (Danaldson and Mavrogordato, 2018, p. 591). For example, in considering and rationalizing which approach to take with a particular low-performing teacher, principals cognitively applied the categories "deserving" versus "undeserving." But once working with a low-performing teacher, relational aspects came to the fore because principals wanted to build trust through demonstrations of benevolence. And when a decision point for removal of a low-performing teacher arose, it was organizational aspects (e.g., union relations) that most influenced principals' behavior. In all these contexts and aspects, principal shaped the evaluation system in its implementation.

Finally among this group of studies, Woulfin, et al., (2016) used frame analysis to investigate how district leaders communicated strategically about evaluation policy to educators

under their supervision. Despite the tendency of the district's frames to stress evaluation's accountability aspects and to advance solutions for school-level implementation issues, the district did not act to enforce its frames or hold school leaders accountable for carrying out particular practices. It instead only regulated the state-enforced procedures. The district left principals to decide how to implement the teacher development components. The predictable result was a compliance mind-set, plus incoherent implementation activities among schools.

Another set of studies seeks to explain how principals alter evaluation demands in their implementation by identifying patterns in principals' goal-oriented actions during the process of interacting with teachers within practical constraints. To take a prominent example, Kraft and Gilmour (2016) reported decidedly mixed results for principals' capacity to effectively use evaluation for formative purposes. On the one hand, the researchers found that evaluation system provided improved materials and positively shifted teacher roles in and the culture around evaluation. On the other, they identified four "challenges" to evaluators that had serious "consequences" for capacity of the system to improve teachers. First of all, it turned out that principals did not agree about the purpose of evaluation, which meant that they used evaluation differently. Second, the expanded evaluation responsibilities of principals led to brief, infrequent discussions of feedback. Third, because principals inevitably had to give feedback outside their expertise, feedback focused on purely pedagogical practices and strategies. Lastly, principals lacked sufficient training; principal-teacher discussions tended to concern summative ratings and positive talk, rather than feedback for improvement (see also Kimball, 2002; Reinhorn and Johnson, 2015 [main challenges to evaluation for teacher development found to be evaluator time and expertise]; Rigby, Larbi-Cherif, Rosenquist, Sharpe, Cobb, P., and Smith, 2017 [most instructional feedback concerned content-neutral classroom practices instead of content-specific

pedagogical practices]; Sporte, Stevens, Healey, Jiang, and Hart, 2013 [evaluation system placed high demands on administrator time and capacity]).

The work of Kraft and Gilmour (2016) is consistent with a larger body of research suggesting that "researchers have found that evaluation is not performed in optimal conditions and is often not effective" (Maslow and Kelley, 2012, p. 602). Maslow and Kelley (2012, pp. 602-603) proceeded to catalogue many of the deficiencies in how evaluators common enact their roles:

> Supervisor competence (or lack thereof), inadequate time for obsevation and feedback, limited teacher/administrator understanding and acceptance of the evaluation tool, and the application of narrow conceptions of teaching impede the ability of evaluators to provide meaningful … feedback to enhance teaching practice [references omitted]. Lack of district-level support in defining criteria by which to evaluate teachers, untrained evaluators, and unclear expectations continue to pervade the field as well [references omitted].
>
> Research also has shown that classroom observations of teachers are often biased and subject to evaluator preferences and "vision" of what effective practice should look like [reference omitted]. In addition, an inherent conflict exists between the role of evaluator as instructional leader and as staff supervisor [reference omitted]. And research has shown that principals lack content-specific knowledge, resulting in evaluation feedback that focuses on general behaviors, such as delivery, rather than on content-specific pedagogy. Thus, evaluation may present an inaccurate view of teacher performance, … [references omitted].

(See also Murphy, et al., 2013; Reinhorn and Johnson, 2015; Woodland and Mazur, 2015).

In a second important study, Kraft and Gilmour (2017) turned to the summative or accountability purpose of evaluation, seeking to explain their demonstration that most principals issue very, very few low effectiveness or underperformance ratings to their teachers. The authors found their explanation in "realities and complex incentives evaluators must navigate when assessing teachers" (Kraft and Gilmour, 2017, p. 240). Principals' "coping and rationalizing behaviors included the accommodating severe time constraints, crediting teachers' upside

potential and positive motivation, avoiding personal discomfort, and minimizing difficulties in removing and replacing teachers" (Kraft and Gilmour, 2017, p. 240). The authors noted that the principals' behaviors are of a piece with those Lipsky (2010 [1980]) attributed to "street level bureaucrats" (see also Weatherley and Lipsky, 1977); and they concluded that while policy may set the general terms of the evaluation system, "it is the aggregation of individual decisions by evaluators navigating complex realities in schools that determine" the distribution of effectiveness ratings (Kraft and Gilmour, 2017, p. 242).

Although their research motivation differed, Donaldson and Woulfin (2018) also wanted to know how principals used their discretion in evaluation, not only to affect outcomes but more tellingly to fashion evaluation processes based on their judgment of what would be the better course for themselves and teachers in the circumstance. Relying on structure-agency theory, they collected and analyzed a large data set of "principals' discretionary activities" from 37 principals and 7 assistant principals from 37 schools in 9 districts then piloting a state-legislated teacher evaluation system. They first showed the ample opportunities for principal discretion among the defined structures of system. The law specified, for example, the number of teacher student learning objectives (SLOs) and classroom observations, but left open the content of the SLOs and the particular classes/lessons observed. Donaldson and Woulfin (2018) then established the substantial breadth of the participants' agency in the day-to-day implementation of evaluation in their schools and with their teachers, particularly with respect to those evaluation structures with which they were most comfortable.

More specifically, they construct a typology of six kinds of principal actions during their conduct of evaluations. The most common, "tinkering," involved making "minor adaptations or adjustments to rules or practices"—for example, an extension in the window for conducting

required midyear conferences (Donaldson and Wolfin, 2018, p. 543). Other strategies included (in descending order of frequency) "reducing," "framing," "hybridizing," "double dipping," and "gaming." The authors also explore the principals' "espoused rationale" for their adaptations of the systems' requirements—from clarifying or simplifying components to enhancing or improving them, increasing buy-in or feasibility, and promoting learning, and reducing anxiety. Principals were trying in the main both to manage the evaluation demands on them and to advance evaluation's developmental goals. Principal agency in institutionalized evaluation seems a fact; whether it helps or hurts the achievement of policy goals in the longer run, and how teachers engage and respond year, after year all remain subjects of future research.

Reinhorn and Johnson (2015) approached the question of how principals' implementation of evaluation reflected differing attitudes and purposes through a multiple case study of six schools of varying grade levels in a single urban district serving a high poverty population. Collecting substantial data from administrators and teachers, the researcher found across all schools that, while teachers accepted, even embraced, both purposes of evaluation, they very much sought high quality feedback for professional growth and school improvement. However, given the attitudes and activities of the school administrators, Reinhorn and Johnson (2015) found that principals actually implemented evaluation to serve quite widely divergent purposes (see also Marsh, et al., 2017; compare Reinhorn, et al., 2017). In four of the six schools, participants described "the evaluation process largely as a ritualized, bureaucratic activity that prompted procedural compliance, but failed to support either increased accountability or professional growth" (Reinhorn and Johnson, 2015, p. 5). The principals in the remaining two schools

> invested heavily in evaluation in order to improve the quality of teaching in their school. However, even these two principals viewed the pur-

> pose of evaluation quite differently. One relied on evaluation to identify and document the shortcomings of teachers who did not meet standards and then to dismiss those who failed to improve. The other used evaluation to achieve both purposes—holding teachers accountable and supporting them for improvement—which teachers widely praised.

In the four schools that engaged in "perfunctory implementation," most considered evaluation "an empty requirement with little or no benefit for their learning." The process focused on procedure; it was a "routine that teachers endured, but did not value" (Reinhorn and Johnson, 2015, pp. 18-19). In the school where the principal evaluated all teachers, but only to remove the weakest, the other teachers expressed relief at the departure of underperforming peers, though another result was a "demoralizing climate of fear." Focusing on accountability for the few, school administration paid little attention to improvement for the many. For most, again, evaluation was largely a "routine, empty, bureaucratic process" (Reinhorn and Johnson, 2015, pp. 21-22). Only in a single school did the principal use evaluation to support teacher growth. He held strict accountability expectations, which unsettled some teachers, but the principal competently dedicated himself to improving teacher quality, despite challenges of time and incomplete content knowledge.

The literature on how and why school leaders conduct evaluations as they do affirms that "[p]rincipals are at the center of implementing these policies" (Donaldson and Mavrogordato, 2018, p. 597). Perhaps given the frontline role principals play in directly executing evaluation demands, it is not surprising that teachers' perceptions of evaluation may to a large extent turn on principals'—but *not* just principals'—behaviors in implementation (Jiang, Sporte, and Luppescu, 2015; Reinhorn and Johnson, 2015; Reinhorn, et al., 2017; Sporte, et al., 2013; Tuytens and Devos, 2010). Jiang, et al. (2015), for one, found that teacher perceptions of professional community were, like school leadership, positively, significantly related to

perceptions of the evaluation system (see also Marsh, et al., 2017; Reinhorn, et al., 2017 [both discussed in the next section of this chapter]).

Speaking more generally, teacher perceptions of evaluation can influence the kinds of responses with which this dissertation is concerned—namely, how teachers engage with the evaluation system, how their participation contributes to evaluative activities (Jiang, et al., 2015; Reid, 2020; Reinhorn and Johnson, 2015; Reinhorn, et al., 2017). The research establishes, for example, that teachers support both the accountability and development goals of evaluation, the former to the extent it accurately identifies underperforming teachers, the latter because the observation process, properly conducted, can result in teacher growth (Donaldson, 2012; Jiang, et al., 2015; Moran, 2017; Sporte, et al., 2013). Teachers unsurprisingly find more value in evaluation as a formative process, rather than for its summative outcomes (Kimball, 2002; Peterson and Comeaux, 1990). They seem to believe that properly applied observation tools align with authentic aspects of their practice and can lead to productive conversations about it (Jiang, et al., 2015; Kimball, 2002; Peterson and Comeaux, 1990). Teachers are particularly concerned about the competence, preparation and fairness of their evaluators (Donaldson, 2012; Kimball, 2002; Reid, 2020; Sporte, et al., 2013). Reinhorn and Johnson (2015, p. 26) sum up the prior points, when they write that their teacher-participants agreed that

> … evaluation has the potential to improve the quality of instruction and, ultimately, student learning. Further, they endorsed the use of evaluations to increase accountability. But they also hoped that evaluation would support their professional growth. Across the sample, teachers longed for opportunities to receive detailed, useful feedback, coupled with support for improving their practice. They criticized or dismissed the evaluation process when it lacked those elements.

As the last sentence quoted above suggests, there is some evidence that evaluation in practice may operate to discourage or diminish the quality of teacher engagement or response.

Some teachers have expressed concerns over the accuracy, fairness, and usefulness of the observation and student growth components (Donaldson, 2012; Hewitt, 2015; Jiang, et al., 2015; Kimball, 2002; Moran, 2017; Sporte, et al., 2013). Another nuance in an assessment of teacher responses is the evidence that teachers' individual characteristics and school contexts—for example, teaching experience, grade level, assignment, prior effectiveness ratings, instructional philosophy—influenced their perceptions and so presumably participatory responses (Jiang, et al., 2015; Kimball, 2002; Lane, 2020; Peterson and Comeaux, 1990). Nevertheless, together with the implementation demand literature, the sparser response literature suggests that in the evaluation implementation dance, for better or worse, teachers do well to follow their principal's lead (see Lane, 2020; Marsh, et al., 2017; Reinhorn and Johnson, 2015; Reinhorn, et al., 2017).

Three recent studies reflect differing perspectives on the principal-teacher dance metaphor (Reid, 2020; Lane, 2020, and Marsh, et al., 2017). Reid's (2020) did not limit his study of teachers' perceptions of how principals use a new evaluation system. He took a further step of showing how those perceptions affected teacher involvement in the evaluation process. Reid (2020, p. 132) points out that:

> … a teacher's perception that his or her principal is utilizing evaluation in ways that may help them improve their practice may help teachers buy-in to the evaluation process. This buy-in has the potential to result in a more meaningful and useful evaluation experience for principals and teachers, as well potentially leading to better implementation efforts desired by policymakers and school districts.

Reid's (2020) analysis of data from nine teachers and nine principals in traditional and charter elementary schools in Michigan resulted in two findings regarding teachers' perceptions: first, their principals' evaluative activities supported the teachers' professional growth, mainly by providing balanced and actionable feedback; and second, their principals acted with transparency

and communicated effectively during the evaluation process. These perceptions together led to teachers responding with active engagement in the evaluation process (Reid, 2020).

Lane's (2020) argument is nearly orthogonal to Reid's (2020). The former applied a frame analysis approach to understanding how teachers responded to evaluation demands based on their beliefs about their evaluators and the purpose of evaluation. He found that teachers in the schools he studied collectively shared an understanding that focused on the summative goal of evaluation. Teachers' accountability-focus conception entailed a "single elaborated perspective" that informed and guided their responses to evaluation demands:

> The primary purpose of evaluation is to earn a high score relative to one's peers and in so doing secure one's occupational future. In order to perform well on evaluation, teaches need to determine a principal's preferences and then project themselves in ways that align with these preferences. Finally, principal preferences are likely to be only loosely connected to the observation protocol, if connected at all, and may actually conflict with a teacher's own sense of best practices.

(Lane, 2020, p. 23; completely italicized emphasis omitted and underscored emphasis supplied).

Probably rooted in teachers' pre-evaluation policy experiences, Lane (2020, p. 19) further found that this perspective motivated teachers' efforts in the evaluation process given their self-understanding of their personal position, irrespective of "differences in principal leadership, local labor market conditions, and teachers' own beliefs, values, experiences (i.e., perceptions), and relative standing among colleagues." These findings provided the foundation for Lane's (2020) further explanations for the effects of evaluation on teacher collegiality and practice.

Lane (2020) pointed out that his approach and analysis differs from that of Donaldson and Woulfin, 2018 and Reinhorn, et al., 2017 (discussed above) because the principals in his study did not impose their evaluation frames on their teachers. Instead, the teachers developed a common evaluation perspective independent of their principals' attitudes, beliefs, or practices,

and they then responded to evaluation accordingly. Still, to the extent that teachers' evaluation responses stemmed from their understanding of the preferences their principals had communicated (as well as their individualized situations), the teachers *were* following or aligning with their principals' leads, at least in significant part. Lane (2020) in addition suggests that, contra Marsh, et al., 2017, "rather than being collectively determined at the school level by a set of organizational conditions, [teacher] responses to evaluation emerged from a common group perspective and one's understanding of her specific situation" (Lane, 2020, pp. 38-39).

As Lane discusses, Marsh, et al. (2017) did focus their investigation on organizational factors and conditions around the implementation of a revamped teacher evaluation system in eight case study schools under various governance arrangements in New Orleans. The researchers did conclude that leadership quality and collaborative structures were the most strongly related to the variation in school-level responses to evaluation demands (discussed in the next section of this chapter). For present purposes, with respect to evaluation responses of teachers and administrators within schools, it is significant that Marsh, et al. (2017) found three non-exclusive patterns of response, which they designated "reflective," "distortive" and "compliant." Importantly, most of the case schools showed a combination of response types based on teacher and administrator response data the researchers describe.

To begin, "[w]hat distinguished distortive schools was that teachers reported changing their behavior during evaluations to enhance their results" (Marsh, et al., 2017, p. 553). The distortive behaviors included "putting on a show" for observations and gaming levels or selecting students to meet student learning targets. As the name suggests, educators in compliant cases "appeared to go through the motions, but did not reflect or act to improve practice" (Marsh, et al., 2017, p, 554). Their responses did not actively "resist, game or adjust their practice to

improve their perceived effectiveness or preclude valid measurement" (Marsh, et al., 2017, p. 554). Notwithstanding relatively minor variations in individual responses, most educators did not respond strongly in either direction to the evaluation components; most reported continuing with their prior practice. A minority of three schools responded in a primarily reflective way. "Educators in these schools reported in engaging in meaningful reflection and improvement efforts, and clearly perceived the evaluation data as a valid measure of teaching and useful for improvement" (Marsh, et al., 2017, p. 556). For example, they committed to the validity and utility of the evaluation process, which was treated as an integral component of their professional practice. Observations were rigorous, but support was strong too, with quality feedback and coaching typically provided. In all the case schools, the agentic implementation of evaluation demands and responses resulted in identifiable patterns of interaction that effected the policy design structures and goals.

Overall, it was and remains unclear the extent to which performance evaluation systems in implementation—especially since they have become increasingly routinized in practice—can meaningfully achieve or consistently sustain either summative accountability or formative development goals (Lavigne, 2014; see also, e.g., Marsh, et al., 2017; Reinhorn and Johnson, 2015), let alone whether they "are capable of reconciling the marriage of teacher development and dismissal in one single system" (Kraft and Gilmour, 2016, p. 740; see also, e.g., Murphy, et al., 2013; Reinhold and Johnson, 2015). As Papay (2012) suggested it would, the summative goal in implementation seems to predominate over the formative (Reinhorn and Johnson, 2015; Woulfin, et al., 2016; see, e.g., Cameron-Standerford, et al., 2017; Ray Taylor and Associates, 2017). The common objective of teachers and administrators often extends little beyond formal compliance (Reinhorn and Johnson, 2015; Woulfin, et al., 2016; Woodland and Mazur, 2015),

notwithstanding the system's substantial opportunity costs (Hallinger, et al., 2014; Murphy, et al., 2013).

Even the evidence on the implementation of each goal considered separately is mixed. With respect to the summative purpose, some research showed increases in low-performing teacher exits with the initial implementation of reforms (e.g., Dee and Wyckoff, 2015; Loeb, et al., 2015). It nonetheless seems that a large majority of teachers in nominally "high stakes" systems receive satisfactory effectiveness ratings (Kraft and Gilmour, 2017; Woodland and Mazur, 2015). Nor do evaluations systems well serve their development purpose, even when school leaders attempt to fulfill it; many studies document the many ways formative demands and responses fall short (Maslow and Kelley, 2012; Kraft and Gilmour, 2016; Woodland and Mazur, 2015; see also, e.g., Rowan, Schilling, Spain, Bhandari, Berger, and Graves, 2013; Serdiouk, Bopp, and Cherasaro, 2017). Still, it is not impossible to implement an effective formative-oriented evaluation system, but formative system elements must be coherently integrated with other organizational supports or cultural conditions for teacher learning, which is challenging and costly for most districts and schools (Marsh, et al. 2017; Maslow and Kelley, 2012; Reinhorn, et al., 2017; Taylor and Tyler, 2012). Opportunities for success with formative evaluation, in sum, may depend on positive interactions in the implementation of otherwise separate reforms.

PLC and Evaluation Implementation Interactions

The literature on the processes and results of interactions between PLC and teacher evaluation in implementation is sparce and diffuse. A handful of studies argues, largely on theoretical grounds, that anticipated conflicts between the two reforms in implementation could impede the achievement of either's goals. Two in this first group go on to propose ways to

mitigate or manage the conflicts. A larger second group comprises empirical investigations of evaluation implementation in contexts where the work or elements of professional community or professional culture contributed to or suffered from implementation results. Finally, some studies of *school level* accountability measures find effects on the goals, work or results of professional community. This arguably analogous research could prove relevant because the identified mechanisms and effects might apply, a fortiori, to the more recent implementation of more intense individual evaluation systems.

Darling-Hammond, et al. (1983) shows in a comprehensive literature review what well know even mor than three and one-half decades ago: teacher evaluation systems can adversely affect elements of school and professional culture. Decades later Darling-Hammond (2013, p. 3) makes a stronger argument in the introduction to her book-length teacher evaluation model:

> …that we not adopt an individualistic, competitive approach to ranking and sorting teachers that undermines the growth of learning communities which will, at the end of the day, do more to support student achievement than dozens of the most elaborate ranking schemes ever could.

Although neither Forsythe, et al. (2011) and Valli, Croninger, and Walters (2007) studied PLCs directly, they both provide examples of how such negative interactions might occur. The former argues that non-evaluative, "supervisory practices" such as shared inquiry, reflective dialogue, collaboration and professional learning, build the collective trust necessary to improve teaching and learning. This collective teacher development approach would only suffer with the adoption an individual performance evaluation system (Forsythe, et al., 2011; Hallinger, et al., 2014). The latter advises wariness toward polices that purport to isolate the contributions of individual teachers to student learning. Even if measurement issues could be resolved, Valli, et al. (2007) maintain that such accountability demands could harm school efforts to encourage teacher

collaboration and more productively distribute resources. They reason that teachers would have scant incentive to spend time together, share resources, or engage in professional activities for other teachers' development.

Two more ambitious studies first identify the same sort of challenge implementing formalized PLCs and performance evaluation together—potential incoherence to the detriment of how both are supposed to improve teacher quality—before arguing for their preferred solution through integrating the systems (Jones, Bettini, and Brownell, 2016; Woodland and Mazur, 2015). Jones, et al. (2016) locates the source of the potential problem in the incompatibility of individual and organizational interests activated by differing accountability mechanisms. They then look to how other occupational fields (e.g., surgery, the military) with the same sort of individual/team accountability tensions address the problem via the explicit evaluation of individual and team collaboration. The authors discuss and propose further research on three constructs by which to determine team performance could be applied to teaching. They conclude that "if the two strands of policy reforms are to coexist, we urge researchers and practitioners alike to take seriously the need to make space in the evaluation process for the assessment of teamwork" (Jones, et al., 2016, p. 8; see also Kimball, 2002).

Woodland and Mazur (2015) compare the "promises and pitfalls" of PLCs with those of teacher evaluation. Not surprisingly, the implementation of both these "hug" and "hammer" approaches, which compete for time, attention and resources, results in an incoherence harmful to teacher professional learning and quality. The authors elaborate solution envisions a wholly new "tiered system of job-embedded professional development (JEPD) that leverages the strengths of both PLCs and educator evaluation as they are currently implemented, minimizes their respective weaknesses, and treats them as a coherent approach" (Woodland and Mazur,

2015, p. 7; emphasis deleted). In the plan's first tier all teachers are members of teams engaged in JEPD and individually subject of administrative observation and feedback. The succeeding tiers mainly ramp up the feedback and other supports for the fewer teachers whose performance continues to lag. Eventually, teachers evaluated as ineffective after the third tier could be dismissed. The system in this way attempts to achieve both the individual and collective professional development to which PLCs aspire, as well as to accomplish the summative and formative ends of evaluation.

Last among those researchers who analyze potential PLC/evaluation interactions in more theoretical terms, Johnson (2015) seeks to demonstrate that the use of value-added measures (VAMs) to identify, reward or dismiss teachers could decrease school capacity for teacher professional community and internal accountability by, as her title has it, "Reinforcing the Walls of the Egg-Crate School." Beyond depressing teacher satisfaction via the creation of an uncongenial work environment, Johnson argues that VAMs would likely discourage teachers' shared responsibility for students and constrain the benefits of standards-based evaluation. Based on social capital theory, Johnson wants to show that the use of VAMs to determine personal accountability would short-circuit the process by which school social capital boosts individual human capital. Rather than relying on VAMs to reward or sanction a few teachers, Johnson (2015, p. 122) concludes that schools would likely gain more through investments in "promoting collaboration, learning, and professional accountability," perhaps including a peer assistance and review process for teacher evaluations.

A small portion of the evaluation literature empirically investigates how the implementation of teacher evaluation systems in particular contexts interacts with collaborative structures and opportunities akin to PLC. In one sub-set of studies, such interactions sometimes

directly yield adverse effects on collaboration. For example, Hewitt (2017) found that the implementation of a particular teacher assessment system utilizing a VAM had several negative effects on teachers, including a decrease in collaboration and an increase in competition among teachers. Stone-Johnson (2015) found that teacher alienation due to accountability demands increased isolation and decreased collegial interactions with counselors over shared student responsibilities.

In a second somewhat more developed sub-set of empirical studies, one issue involved the role collaborative structures play upon the implementation of a teacher evaluation system. The variance in findings on this issue is shown in the results from several research projects. First, as a part of a larger, multifaceted endeavor, Reinhorn, et al. (2017) reported on how six high-performing, high-poverty schools implemented the then-new Massachusetts multiple-measure teacher evaluation (MMTES) system (see also Reinhorn and Johnson, 2015; Johnson, et al., 2018). Their sample included all types of public schools (traditional, turnaround, restart, and charter), all of which had received the state's highest accountability rating. The researchers sought to ascertain how these successful schools approached putting the state-mandated evaluation system into practice. Two findings highly relevant to evaluation/PLC interactions stand out from their detailed analysis. First, the schools integrated evaluation with other professional learning opportunities, including a form of PLC, as one component in a coordinated strategy for teacher improvement.

> In explaining the support they received, teachers often did not distinguish between practices that were part of the evaluation system and others intended to improve their practice; as they say it, all were part of an ongoing, integrated improvement process.

(Reinhorn, et al., 2017, p. 398). The quotation points to a second finding about how evaluation could be successfully implemented to improve teacher practice: the schools did not use

evaluation primarily for summative accountability; rather, they focused on the formative potential of evaluation. As the authors point out, "many [teachers] identified classroom observations and feedback as the most valuable component of their school's developmental process" (Reinhorn, et al., 2017, p. 398; see also Taylor and Tyler, 2012).

In a second comprehensive study that, like Reinhorn, et al. (2017), examined the implementation of a new MMTES, the sample in this case being eight schools in the New Orleans recovery district (Marsh, et al., 2017). The researchers identified three types of school responses to evaluation that they denominated "reflective," "distortive," and "compliant." They considered the potential contribution of a number of factors to the schools' responses to evaluation. They found that two school level organizational factors relating to teacher professional community facilitated the more beneficial "reflective" response—namely, shared instructional leadership, and structures for frequent collaboration. Among other things, the latter factor led to authentic collaboration with a "sense of accountability for continuous improvement[,]" and "also provided an avenue for enhanced peer accountability" (Marsh, et al., 2017, p. 561).

Third, Tichnor-Wagner, Harrison, and Cohen-Vogel (2016) conducted research on four large comprehensive high schools, two of which were highly effective and two less effective, to analyze the elements of a "culture of learning" present in each. The authors engage in an extended discussion of those elements, and in particular attend to the role of professional community in the face of research showing the threats posed by teacher evaluation and control-oriented PLCs. Nonetheless, Tichnor-Wagner, et al. (2016) found that the cultures of leaning in the effective high schools were not only stronger in general, but also had developed specific structures and practices, as compared to the less effective high schools. Further, these key

elements included PLCs as structures that enhance formal collaboration; indeed, "key differences between [the two types of schools] were found in the mode and structures of collaboration and the frequency with which collaborative structures were enforced" (Tichnor-Wagner, 2016, p. 619).

Several researchers took varying approaches to framing the issue of evaluation/PLC relationships. Maslow and Kelley (2012) asked under what conditions teacher evaluation provides meaningful learning. They found that teacher evaluation could promote teacher professional learning through meaningful feedback. However, such evaluation required four conditions, two of which were an "organizational culture focused on advancing student learning" and "a strong collaborative school culture with a shared vision of high levels of learning for all students …" (Maslow and Kelley, 2012, p. 628). Jiang, et al. (2015) relied on a dataset from a large, detailed survey of Chicago teachers' perceptions of the then-new teacher evaluation system. These authors concluded that teachers' views on evaluation varied with their perceptions of the professional community leadership at their respective schools (teacher attributes played a role as well). Richmond and Manokore (2010) took a different approach: they introduced voluntary elementary grade-level PLCs as a targeted intervention to improve teachers' elementary science instruction. Although the authors succeeded in establishing functional science PLCs, accountability pressure mainly affected science instruction indirectly. The emphasis on mathematics and literacy crowed out the time and attention teachers could devote to science. They maintained nonetheless that their PLC experiences helped teachers adjust to these constraints.

A last group of research may provide indirect or analogic evidence bearing on the interaction of teacher evaluation and PLCs. These studied how the implementation of *school*

level accountability reforms acted upon PLCs, or at least the defining elements of professional community. Thus, from her studies of probationary schools in Chicago, O'Day (2002, p. 308) observed that the more rapidly improving schools probably "differed significantly from other probation schools along several dimensions of initial school capacity: peer collaboration, teacher-teacher trust, and collective responsibility for student learning [reference omitted]." She later concluded that "[o]ur research in Chicago suggests that bureaucratic school accountability policies are insufficient to establish the patterns of interaction that might foster more effective information sharing in low-capacity schools" (O'Day, 2002, p. 311; see O'Day, 2008).

Another study applied symbolic interaction theory to find uniquely negative effects of external accountability on teacher professional community (White and Rosenbaum, 2008). These authors showed that external accountability reoriented professional norms through its effects on prestigious teaching awards, teacher organizational status and control of resources, and teacher organizational stigma and withdrawal of resources. Those teachers who changed their professional attitudes and behaviors to obtain awards, and used strategies to achieve accountability goals, such as meeting test score targets gained status and access to school resources, which were denied other teachers who could not compete with them. The result was that, as external accountability distorted individual teacher behavior, it likewise altered school culture. Talbert and McLaughlin (2008) have similarly argued that external accountability encourages traditional teacher norms, such as private practice and individual autonomy, thereby undermining social cohesion and sustained collaboration.

In contrast, two researchers have found largely positive effects. Using extensive data from 32 schools in ten districts across five states, Desimone (2013) demonstrated that standards-based external accountability engendered positive change in several areas related to teacher

professional community and accountability, such as attention to struggling learners, responsibility for student learning, and classroom content and pedagogy. While acknowledging the stress and pressure associated with the reforms, Desimone (2013, p. 33) concluded that "[t]he findings here show that the testing and accountability system had moved schools in the desired direction—toward personal and group responsibility for student learning. Desimone further found that teachers had assumed substantial responsibility for responding to the reforms and changing their practices. She suggested that the transparent nature of the accountability (reported test results) encouraged teacher professionalism.

Weathers' (2014) reached conclusions like Desimone's, but via a very different methodology. He set out to investigate the roles played by principal leadership and external accountability in teachers' understanding of professional community. He used multi-level modeling of large-scale survey data to estimate the effect of school-level performance accountability policies on teacher community. His findings ranged from statistically non-significant to weakly significant, depending on model specifications. He suggested that "top-down policies that encourage (re-)examination of the core practices may foster a sense of community" as teachers work on a common set of practices to develop "a common language, trust, and mutual respect and accountability" (Weathers, 2011, p. 23).

Another a small group of studies primarily concerns the effect of external school-level accountability on individual teacher attributes, from which a line is drawn to determinants of teacher community relating to professional or internal accountability. Buchanan (2015), for instance, studied the interaction of external accountability demands with teacher identity and agency; teacher characteristics and local context were considered mediating factors. She concluded that although teachers sacrificed professional identity and professionalism (in terms of

autonomy) due to external accountability pressure, resistance to these losses was most successful in schools with strong collaborative communities sharing a collective set of commitments and practices. Banerjee, et al. (2017) found that a "collective pedagogical teacher culture," including a professional community of trust and common focus, as well as collaboration with collective responsibility, buffered the negative effect on student achievement of threats to teacher satisfaction level. However, Erichsen and Reynolds (2020) found nearly the opposite—i.e., collective pedagogical teacher culture, which they conceptualized as professional culture and teacher collaboration, did *not* buffer the otherwise deleterious effects of accountability pressures on teacher morale and job satisfaction (see also Finnigan and Gross, 2007). Another study involving teacher satisfaction and school organization under external accountability pressure highlighted the significance of professional and internal accountability (Johnson, et al., 2012). The researchers there investigated the work conditions most contributing to teacher satisfaction. The most significant were collegial relationships, principal leadership, and school culture (including trust and commitment). They then showed that the self-same elements were the strongest determinants of student achievement growth, and hence of school response to external accountability. Finally, and more generally, in Valli and Buese's (2007) study of how teacher roles changed under accountability pressures, the collaborative role (like others) expanded and intensified at the same time it became more bureaucratic and less authentic.

Whether positive or negative, direct or indirect, there's *some* relationship, and often interaction, between professional community or culture and school or individual external accountability. Moreover, except where PLC and evaluation are implemented in an intentionally integrated manner, the direction of principal effects seems to run from evaluation to teacher collaboration and community.

Conclusion

The largely siloed development of research pertaining to evaluation and PLC as divergent teacher quality policies seems mainly due to when and how each appealed to policymakers, practitioners, and researchers. A large majority of the relevant implementation literature focuses on evaluation *or* PLC, rather than on any significant relationship or interaction between them. Likewise, although researchers have studied teacher responses to one or the other, we know little about how teacher response to each may change in the face of the concurrent, potentially related demands of both.

Even that research finding or reporting interactions, or potential interactions, fails to bridge the research gap because little or none explicitly examines evaluation and PLCs together as institutionalized or routinized, let alone as formal structures with which teacher must re-engage each school year. Instead, these studies mostly involve (a) some combination of proposals to mitigate or resolve anticipated conflicts between evaluation and PLCs; (b) a focus on the implementation of evaluation, while treating PLCs, or some element necessary to them, as a supportive or obstructive condition or factor; and/or (c) limited or no attention to any mechanism of interaction and effect between evaluation and PLCs. At best, this research suggests that accountability pressures—whether applied at the school or teacher level—may affect or be affected by PLCs or related structures, processes, or resources. With the widespread establishment and operation of evaluation systems and PLC models together, it seems likely that *some* kind of interaction would be inevitable with consequent implications for the effectiveness of each policy, as well as for the coherence of both.

A substantial gap appears in the literature, one precisely where research would appear especially salient and beneficial to policymakers, practitioners and future researchers. This dissertation seeks to address that gap.

CHAPTER 3
POLICY CONTEXT AND CONCEPTUAL FRAMEWORK

The wide spread adoption and implementation of teacher evaluation systems and PLC models has occurred in a wide range of contexts. As detailed in Chapter 4, this study concerns these policies as institutionalized in Michigan districts and schools, and therefore in the Michigan evaluation and PLC policy context, the subject of the first section of this chapter. It describes the evaluation system mandated by state law, including the discretion left to districts and schools. The first section will likewise summarize the voluntary adoption by many Michigan districts of PLC models that frequently entailed formalized grade level/disciplinary content structures, goals and practices. In light of this context, the second section of this chapter sets forth the structure-agency lens through which this study examines the phenomenon of interest: teacher interacting responses to the related demands of PLC and evaluation in their schools.

Michigan Teacher Evaluation and PLC Policy Context

Michigan joined the state curriculum standards and testing movement in the 1990s that led to the adoption of NCLB (Mehta, 2014). Due mainly to the Obama Administration's NCLB waiver policy and its Race to the Top (RTTT) program, teacher evaluation shot to the top of the policy agenda in many states, including Michigan. As part of its (unsuccessful) 2009 RTTT application, Michigan rapidly enacted the first prescriptively detailed teacher evaluation system as amendments to its Revised School Code (PA 2009, No. 205, Imd. Eff. Jan. 4, 2010, codified as Mich. Comp. Laws §380.1249; see PA 2010, No. 336, Imd. Eff. Dec. 21, 2010). Other enabling legislation amended (1) the Public Employment Relations Act, Mich. Comp. Laws §§423.201, et seq. by prohibiting bargaining over teacher evaluation and related management rights, such as discipline and discharge, and lay-off and recall, and (2) the Teacher Tenure Act, Mich. Comp. Laws §§38.71, et seq. by substantially weakening substantive and procedural

tenure protections (Pogodzinski, et al., 2015). The opening sentence of the evaluation statute mandates that local district boards of education "shall adopt and implement for all teachers and school administrators a rigorous, transparent, and fair performance evaluation system that does all of the following: …" (Mich. Comp. Laws §380.1249(1)). Substantial revisions to the system occurred in 2011 (PA 2011, No. 102, Imd. Eff. July 19, 2011); and further, based on additional study (Michigan Council for Educator Effectiveness [MCEE], 2013), the statute was more extensively updated in 2015 (PA 2015, No. 173, Imd. Eff. Nov. 5, 2015; see Michigan Department of Education, nd.b; Pogodzinski, et al., 2015). The policy turbulence has not completely abated: the most recent amendment, which again changed the weight given "student growth and assessment" measures in teachers' final ratings, was enacted in 2018 for district and school implementation beginning in the 2018-19 school year, with an additional change beginning in the 2019-20 school year (PA 2018, No. 235, Eff. Sept. 25, 2018).

According to the Michigan Department of Education (MDE), "[i]mplementing Michigan's educator evaluation law with fidelity is a key strategy in our [MDE] efforts to see Michigan become a top ten education system within the next ten years." Moreover, the proper conduct and use of individual teacher evaluations form an essential support for the state's strategic goal "to 'develop, support, and sustain a high-quality, prepared, and collaborative education workforce'" (MDE, nd.b, p. 4). Michigan's teacher evaluation system is supposed to serve both summative and formative purposes (Mich. Comp. Laws §380.1249; MDE, nd.b; Serdiouk, et al., 2017).

The evaluation system applicable to the school year of this research (2019-2020) required an annual, year-end evaluation for all teachers that included the following:

- Forty percent of the evaluation must consist of student growth and assessment data; for core content areas (subjects and grades) subject to state assessment 50% of student growth must be measured by state assessment data; in other subjects and grades, this component must be based on multiple, rigorous, research-based and district-wide measures.
- The other 60% of a teacher's evaluation must be based primarily on their performance as measured by an observation tool, which is applied to a minimum of two classroom observations (one unscheduled) during the year. (First year probationary teachers and those previously rated low-performing are subject to additional performance planning and reporting.) The observation tool, incorporating a research-established teacher performance framework/rubric of teaching standards, must be state approved and district adopted.
- The district must provide teachers with designated student data. Within 30 days of each classroom observation, teachers must also be provided with feedback from an observation. Teachers and evaluators must consult to develop performance goals and recommend professional development for the following school year. These requirements are intended as formative assessment to ensure that evaluation helps teachers improve professionally.
- The summative, ranking purpose of the evaluation system is this: based on the determinations of the evaluation components, each teacher is assigned a numerical effectiveness rating, which falls within one categorical range of "ineffective," "minimally effective," "effective," or "highly effective." A teacher's only appeal lies with the district's superintendent, who is the final decisionmaker.

(Mich. Comp. Laws §380.1249; MDE, nd.b).

A district is supposed to use a teacher's effectiveness rating as the primary determinant of all significant employment decisions concerning that teacher, including renewal/non-renewal of probationary contracts, grants of tenure status, and lay-off and recall (Mich. Comp. Laws §380.1249(1)(d); see Mich. Comp. Laws §380.1248). Without limiting a district's power to dismiss any teacher for poor performance, the statute explicitly requires dismissal on three consecutive years of an "ineffective" rating (Mich. Comp. Laws §380.1249(2)(j)). A teacher can only have the annual evaluation requirement reduced to a biannual one by receiving three consecutive "highly effective" evaluations and maintaining that rating thereafter (Mich. Comp. Laws §380.1249(2)(k)).

The state statute does provide for district discretion within certain components of the evaluation system (Pogodzinski, et al., 2015). Procedurally, each district selects (and provides the resources for) the technological hard- and software necessary to conducting evaluations each year, along with the system for data entry, retention, analysis and reporting. Each district provides for evaluator training and supervision, even as it must also inform teachers about the process components, timeline and content requirements. More substantively, unless it bears the expense of creating and gaining MDE approval of its own "evaluation tool," each district selects from a state-approved list of five the one tool, the instrument with a specified set of teaching standards and scores, that its administrators will use during classroom observations and that will apply to other teacher evidentiary submissions (Mich. Comp. Laws §380.1249(2)(f)). Each district possesses some discretion to decide on the allowable evidence of student growth, as well as the method by which student growth is ascertained. However, as noted above, the statute's several amendments constrain the scope of district decision making by increasing the weight of

the student growth component and by more narrowly specifying the nature and source of student growth data that must be used (Mich. Comp. Laws §380.1249(2)(a); MDE, nd.b).

Research on Michigan's dual-purpose evaluation system, almost all of it conducted or commissioned by MDE, showed substantial variation in its implementation and in educators' perceptions. To begin with, plausibly due to the obvious "tension between the dual roles of the educator evaluation system" (Ray Taylor and Associates, 2017, p. 6), 97% of surveyed teachers and 96% of administrators surveyed said evaluation *should* be used for a formative purpose, but the respective percentages drop to 42% and 67% when asked for what purpose the evaluation system *is* used. Teacher and administrator perceptions similarly differed with respect to the positive versus negative effects of evaluation on stress, competition, collaboration, student learning, and fairness in identifying under-performing teachers (Cameron-Standerford, et al., 2017).

The reported survey results find support in the evidence of the decidedly mixed implementation of the evaluation feedback requirement, at least from the teacher perspective (Cameron-Standerford, et al., 2017; Marzano Research, 2017; Ray Taylor and Associates, 2017; Rowan, et al., 2013; Serdiouk, et al., 2017). Ray Taylor and Associates (2017, p. 6), for instance, found that teachers "report little useful feedback to improve their professional practice and support their professional growth." Hence, at least on the teacher view, the summative goal seemed to predominate in practice (Cameron-Standerford, et al., 2017; Ray Taylor and Associates, 2017). Further, with the focus on completing the annual summative assessment, the quality, value and utility of the formative assessment components varied widely due to the predictable, understandable, practical constraints on evaluator time, motivation and expertise

(Camerton-Standerford, et al., 2017; Marzano Research, 2017; Ray Taylor and Associates, 2017; Rowan, et al., 2013; Serdiouk, et al., 2017; see Papay, 2012; Woodland and Mazur, 2015).

If the system's formative goal often receives short shrift, the summative goal is far less successful at distinguishing among levels of teacher quality. MDE statewide data shows the percentage distribution of teacher performance ratings between the 2014-15 and the 2018-19 school years:

Table 3.1 Michigan Teacher Effectiveness Ratings

School Year	% Highly Effective	% Effective	% Minimally Effective	% Ineffective
2014-15	42	56	2	0
2015-16	42	56	2	0
2016-17	39	59	2	0
2017-18	40	58	2	0
2018-19	41	58	1	0

Source: Adapted from MI School Data, www.mischooldata.org

In the same period, the total number of evaluated K-12 teachers in Michigan rose from 94,972 to 102,437, a 7.9% increase. In short, while 98-99% of Michigan teachers have been rated effective and above, almost no one has been rated minimally effective or ineffective rate (see Lenhoff, 2012; MDE, nd.a; see Kraft and Gilmour, 2017). If Michigan's evaluation system is a "high-stakes" policy on paper, in practice it is anything but.

Years before the legislature began reforming teacher evaluation state-wide, Michigan enacted reforms in state curriculum standards and assessments (Mehta, 2014). With the advent of NCLB, Michigan joined its sister states to greatly elaborate and strengthen its system of test-based school accountability. Responding in part to school improvement pressures, and in part to the latest reform bandwagon, many Michigan districts and schools, like many nationwide, adopted formalized professional community or teacher collaborative team models (Blankenship and Ruona, 2007; Brouwer, et al., 2012; Charner-Laird, et al.; Feger and Arruda, 2008; Levine, 2010). Despite a plethora of names and configurations (Levine, 2010; Stoll, et al., 2006), by this

time in the reform movement, the formalized PLC models promoted, adopted and implemented share a common theory of action and similar design features, including certain organizational structures and supports, that are considered key to their effectiveness (Brouwer, et al., 2012; Hord, 1997; Levine, 2010; Stoll, et al., 2006).

One commonly implemented structural unit is the teacher team organized and functioning as a grade-level/content PLC. The grade-level/content PLC is an essential building block because so much PLC work involves the curriculum, assessment, instruction, and student learning pertaining to teachers' particular classes, whether organized at the grade level (as in elementary schools) or by department/disciplinary content (as in middle and high schools) (e.g., Dufour, 2004; Szczesiul and Huizenga, 2015). Furthermore, the organization/membership of PLCs is just one structural feature that serves to formalized and routinize PLCs as implemented. There are many others, including, for example, PLC-specific professional development for teachers and administrators; a regular calendar of PLC meetings (which often means altering the school day schedule); specification of PLC goals, tasks, activities and products, which often involving the development of group norms and the use of protocols; supervisory requirements or expectations, typically instantiated in the submission of PLC agendas, minutes, and products, as well as administrator observation of and participation in meetings (e.g., DuFour, 2004; DuFour and Eaker, 1998; Feger and Arruda, 2008; Leine, 2010; Stoll, et al., 2006).

The significance of formalized structures, together with an emphasis on grade-level/content PLC, is exemplified by one of the more prominent models a number of Michigan districts have adopted. This is the PLC model developed by Richard DuFour and Robert Eaker, and widely promoted by DuFour, later with his wife Rebecca DuFour (e.g., DuFour and Eaker, 1999; DuFour et al., 2004). For example, in the early aughts, Oakland Schools, the intermediate

school district (ISD) for Oakland County, Michigan, organized at least three cohorts of constituent districts to receive extensive DuFour training and follow-up support services. At least in southeastern Michigan, leading ISDs, such as Wayne Count RESA and Macomb ISD, began (and continue) providing professional development on the DuFour model. The DuFours and colleagues have continued to present summer PLC "institutes" in Michigan and other midwestern states, and they have created an application and approval process for "Model PLC Schools."

The relevant policy context, in sum, encompasses two formalized, well-established and functioning teacher quality reform policies. In Michigan, state law has imposed a teacher performance evaluation system on districts and schools for nearly a decade. For years, too, many of the same districts and schools have adopted and implemented a PLC model that teachers enact in grade-level/content groups as tasked and supervised by district and building administrators. These reforms were independently developed, adopted, and implemented, even as each made (and makes) its own demands for teacher response (Woodland and Mazur, 2015; Jones, et al., 2016). When multiple policies are in fact implemented in such a context, much uncertainty and many questions necessarily arise around the policies' interaction and joint influences on teachers' work and working conditions, which ultimately affect teaching and learning (Knapp, et al., 1998). Here an apparent fragmentation of reform demands (Srinivasan and Archer, 2018) makes coherence crafting in response a highly important and salient research topic (Honig and Hatch, 2004; Stosich, 2018).

A Structure-Agency Approach to a Coherence Challenge and Resulting Research Questions

Woodland and Mazur (2015, p. 20) aptly point out that:

> American public education is experiencing a watershed moment. Educator Evaluation and PLCs, what we have called "the hammer

> and the hug," are two of our nation's most predominant approaches to instructional improvement, yet in most states Ed Eval and PLCs are not linked, educator account ability and professional learning are widely treated as separate policy initiatives [reference omitted]. These two prevailing reform strategies often appear to exist at cross-purposes….

(See Darling-Hammond, 2013; Johnson, 2015; Jones, et al., 2016; Valli, et al., 2007). Talbert (2010, p. 560) argues more pointedly thus:

> Centralized accountability systems can work against the development of mutual teacher accountability. … Further, emphasis on individual teacher quality, curriculum implementation in each classroom, and monitoring of teaching fidelity undermines principles of collective responsibility and improvement. Teachers' attention is focused on compliance rather than accountability to colleagues for developing new instructional interventions…

The prima facie incoherence of evaluation and PLC, and the problematic position in which teachers find themselves with the implementation of both (e.g., Hewitt, 2015; Stone-Johnson, 2016), are suggested by the comparison indicated in Table 3.2.

Table 3.2 Idealized PLC Model and Teacher Evaluation System Characteristics

Characteristic	**PLCs**	**Teacher Performance Evaluation**
Problem to Address	Teacher isolation, privatization of practice, norm of equality; constrained professional learning and social capital formation	Lack of effective personal performance evaluation and accountability; the "widget effect"
Period of expanding adoption and principal triggering event	c. 1995-2010 NCLB (2001)	c. 2006-present RTTT (2009)
High Level Goal	Improve individual and collective (school-level) teacher capacity/practice	Improve individual and collective (school-level) teacher capacity/practice
Goal, re: teacher professional development	Increase teacher and organizational learning and social capital, applying shared expertise to individual practice	Facilitate improvements in teacher practice through standards-based formative assessment and feedback
Students directly benefited	All students of all PLC teacher-members (common disciplinary content +/or grade level)	Students of each evaluated teacher
Policy design approach	Build professional commitment	Assert bureaucratic control
Primary target/focus	Teachers-members in professional communities (formalized groups)	Individual teachers

Table 3.2 (cont'd)		
Degree of centralization of accountability	Low	High
Level on which external accountability imposed	School (indirectly)	Individual teacher (directly)
Mechanism for individual teacher improvement	Professional development through collaborative learning, teacher commitment, and shared responsibility	Appropriate and actionable feedback based on teaching standards and observation
Mechanism for collective (school) improvement	Improve and sustain organizational learning; align individual and collective teacher responsibilities to improve practice	Use summative teacher ratings to dismiss in- effective teachers and retain/compensate more effective teachers, thus raising the aggregate school level of teacher effectiveness
School-level implementation approach	Bureaucratic to professional	Summative to formative
Primary district role	Decide on adoption of PLC model, provide resources for implementation at school level	Comply with state mandates provide resources for implementation, make/approve teacher employment decisions
Primary building administration role	Implement, supervise and maintain district-selected PLC system, and facilitate desired PLC work and outputs	Comply with state and district mandates to implement evaluation system by conducting evaluations and producing required system outputs
Primary teacher role	Engage with system as peer participant and learner in required tasks to produce desired group outputs and improve practice	Engage with system as subject of evaluation to demonstrate accountability, utilize formative feedback, and obtain summative rating
Teacher Motivation	Primarily intrinsic	Primarily extrinsic
Teacher Peer Relationship	Cooperative/collegial	Comparative/competitive
Individual teacher stakes	Low	High

Source: Dissertation author, as derived from literature.

Teachers, the common policy target of these disparate reforms, at a minimum must annually respond to the demands of both PLC and evaluation, including separately mandated engagements year after year, *both* as individual employees *and* as group members working in nested contexts of the individual, PLC, school, district and beyond (McLaughlin and Talbert, 2001). Furthermore, not only have dual demands required that teachers somehow respond, but over a number of school years they have actually done so, though we know little about how their

responses have interacted or how the interaction of demands as implemented has affected responses.

These very facts nonetheless imply that, despite the potential for conflict between formalized structures of evaluation and professional community, teachers might well develop a strong motivation and adaptive ability to play a role in crafting their own individual and collective coherence, which the concurrent implementation by administration might help or hinder (Honig and Hatch, 2004; see Knapp, et al., 1998; Shirrell, 2016; Stosich, 2018; Tichnor-Wagner, et al., 2016). Such a hypothesis seems plausible for several reasons. First, in the conceptualization of Honig and Hatch (2004), districts and schools mainly craft coherence along a continuum of bridging and buffering actions (see DiPaola and Tschannen-Moran, 2005). A dynamic bridging/buffering strategy could explicitly or implicitly provide linkages, or opportunities for linkages, between evaluation and PLC. Although they do not use the terminology of crafting coherence, the literature on PLC/evaluation interaction includes proposed integration or reconciliation of the two—for example, either through a tiered system of job-embedded professional development based on performance standards (Woodland and Mazur, 2015), or by developing assessments of teacher team performance as a significant component of evaluation (Jones, et al., 2016).

Second, as pointed out in the policy context discussion above, evaluation and PLC elements and products are concerned with many of the same educational content and tasks. The evaluation system in Michigan has components requiring (1) professional learning or growth goal-setting, assessment and reflection; (2) observation, assessment and feedback around curricular standards, lesson planning, instructional practices, and student assessments; and (3) the data collection, analysis and measurement of student performance. Much the same is fodder for

PLC work, albeit occurring in a very different organizational context through very different processes. DuFour (2004, p. 10) at once captures some of what makes PLCs distinctive and what overlaps with evaluation (where emphasis supplied):

> Collaborative conversations call on [PLC] team members to make public what has traditionally been private—*goals, strategies, materials, pacing, questions, concerns, and results.* …
>
> Teams must focus their efforts on crucial questions related to learning and generate products that reflect that focus, such as *lists of essential outcomes, different kinds of assessments, analyses of student achievement, and strategies for improving results.* Teachers must develop norms or protocols to clarify expectations regarding roles, responsibilities, and relationships among team members.

(See also Feger and Arruda, 2008). Practical connections like these could in effect incentivize teachers (and administrators) to locate efficiencies in their engagement with both.

Finally, some research finds that teachers have engaged in coherence-making when two conflicting education policies make concurrent demands on teachers. For example, applying a sensemaking framework, Russell and Bray (2013) studied the responses of special education teachers to perceived inconsistences between NCLB and the Individuals with Disabilities Act (IDEA) that adversely affected their work with students. Similarly, Stosich (2018) in a multiple case study investigated the variance in how teachers and principals in several elementary schools tried to integrate the demands of teacher evaluation and the Common Core State Standards. Robinson (2012) focused on how teachers in Australia developed individual and collective "professional agency" as the means to resolve compliance with two conflicting "policy texts" consisting of student reporting regulations, one issued by the state government, the other by the federal government. Despite differing methodologies, contexts, and mechanisms, all these studies recognize that teachers' responses reflect attempts to address incoherence in converging policy demands.

These reasons, together with the policy context, suggest that concurrent PLC and evaluation demands present a recurring coherence challenge to teachers who are subject to both in a given district/school context. To examine and understand the nature of the challenge and how (if at all) teachers respond to it, this study takes a structure-agency approach. A structure-agency conceptual framework well fits the problem and purpose of this research because, at base, PLC and evaluation demands are structures that condition but do not determine teacher responses; to the contrary, the exercise of some degree and kind of agency is part and parcel of teacher responses. Moreover, through its focus on the *relationship between and the interaction of* policy structure and policy target agency, this analytical approach illuminates policy implementation and institutional change (Burch, 2007; Coburn, 2016; Rigby, et al., 2016). Indeed, "the question of policy implementation is fundamentally about the relationship between social structure and agency"; and further, according to scholars in the field, "the process of policy implementation is, at root, a process of institutionalization [reference omitted]" (Coburn, 2016, pp. 466, 469; see Burch, 2007).

"Structure" comprises the rules, relationships and resources that together constitute regularized patterns serving organizational purposes and functions, such as guiding and constraining organizational activities and behavior, and dividing organizational work and authority (Chingara and Heystek, 2019; Coburn, 2016; Donaldson and Woulfin, 2018; Rigby, et al., 2016). Importantly, structural rules may be formal or informal, and they include social and organizational norms rules and norms (Woulfin, 2016; see Coburn, 2016). From an organizational actor's point of view, structures are how one "understands how things should be done, practices organized around those understandings and capabilities that support those

understandings" (Rigby, et al., 2016, p. 296; see Chingara and Heystek, 2019; Donaldson and Woulfin, 2018).

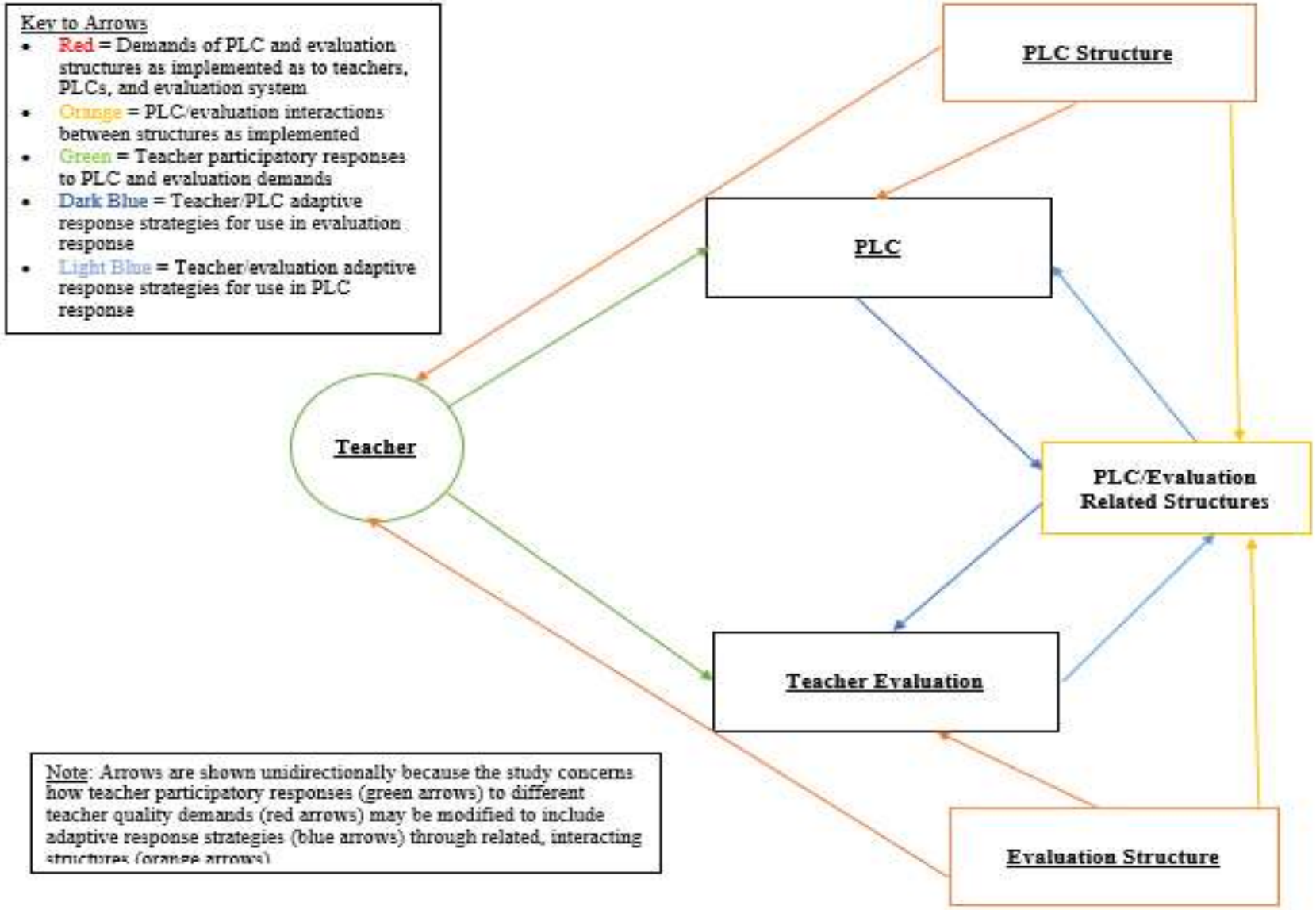
"Agency," in contrast, is an individual's capacity to take action, to act with some social effect (Coburn, 2016; Donaldson and Woulfin, 2018; Rigby, et al., 2016). Agentic capacity includes the ability to alter the constituents of structure—i.e., rules, relationships, resources (Coburn, 2016). The actual use of that ability is shown by actors' discretion in implementation, discretion that is exercised in response to education policy demands (Donaldson and Woulfin, 2018).

The research reviewed in Chapter 2, together with the discussion of the policy context and coherence challenge in this chapter show beyond doubt that PLC and evaluation *each* entail distinct structures, structures that make demands of teachers, including demands for teacher engagement with those very rules, relationships and resources that inform, guide and constrain how "PLC'ing" and teacher evaluating are done (Chingara and Heystek, 2019; Coburn, 2016; Donaldson and Woulfin, 2018; Rigby, et al., 2016). Particularly as formally established and supervised by district/school administration, the structure of PLC includes the organization and membership of teacher teams; all manner of logistics for PLC meetings and work; PLC norms, expectations and oversight; and PLC goals, activities, tasks and products (e.g., DuFour, 2004; DuFour and Eaker, 1998; Feger and Arruda, 2008; Leine, 2010; Stoll, et al., 2006). By contrast, state law instantiates the basic structure of teacher evaluation by mandating that every district have a performance evaluation system meeting certain content and process requirements (Mich. Comp. Laws §380.1249(1); MDE, nd.b). The legally required components, shaped by district/school implementation, involve teacher/evaluator communications and interactions, formal and informal observations, professional standards and assessment rubrics, data collection

and student growth computation, and a final summative rating (Mich. Comp. Laws §380,1249; MDE nd.b; Pogodzinski, et al., 2015; see, e.g., Donaldson and Woulfin, 2018). The structures of PLC and evaluation surely constrain teacher responses to their respective demands, but those responses are not wholly determined. Indeed, by the very nature and terms of the demands, teachers retain and exercise a degree of agency in *how* they engage with and respond to each (e.g., Lane, 2020; Stosich, 2017; Szczesiul and Huizenga, 2015; Stoll, et al., 2006; Thiel, Schweizer, and Bellmann, 2017).

This study extends a structure-agency approach to schools where *two* policy structures with interaction opportunities converge in recursively demanding teacher agentic responses, which themselves may interact through the structures. Figure 3.1 depicts conceptually the relevant structures, agentic responses, and relationships among them.

Figure 3.1 Teacher Responses to PLC and Evaluation Structures with Adaptive Responses Through Related Structures



The red rectangles in Figure 3.1 represent the separate structures of PLC and evaluation as implemented by administration in a school. As the red arrows indicate, those structures each make demands on teachers, both individually and as participants in a particular PLC and with a particular evaluator in the evaluation system. Each teacher-PLC member directly responds to the separate PLC and evaluation demands by agentic participation in each structural process during a school year (green arrows). Teacher agency is not exhausted by the direct participatory responses to PLC and evaluation (green arrows). Rather, it may be expected that, whether by administration implementation design or teacher implementation adaptation, the interaction of the different structures, shown by the orange arrows, may result from related structural elements, which the orange rectangle represents. Now, through the relationship and interaction of PLC/evaluation structures, teacher agentic engagement in PLC can itself affect teacher response to evaluation (dark blue arrows), and vice versa (light blue arrows), in either case, phenomena I will call "adaptive response strategies."

From a structure-agency perspective, it is the dynamic intertwining of structural demands and agentic responses that tends toward aligning the implementations of, as well as altering the outcomes of, both teacher quality policies. A structure-agency framework thus provides a path toward understanding and explaining the phenomenon of interest. It is powerful lens through which to examine *how* teachers, through the exercise of their individual and collective agency in interaction with institutionalized teacher quality structures, participate in the crafting of coherence between them.

Following from the structure-agency representation in Figure 3.1, the statement of this study's research questions rely on structure-agency concepts, as follows:

1. How, if at all, do teachers adapt their agentic responses to performance evaluation and professional learning community demands as they respond to the structures of both?

2. How, if at all, does the relationship between structures of performance evaluation and professional learning community shape teachers' adaptive responses?

The first research question focuses on the teacher agency side, the second on the policy structure side of the structure-agency approach. But the questions together require an investigation into the relationship between policy demand structures of and teacher agentic responses to evaluation and PLC. Accordingly, the next chapter addresses the research design and methodology that is well-fit for this research task.

CHAPTER 4
RESEARCH DESIGN and METHODOLOGY

To answer its research questions, this dissertation relies on a multiple explanatory case study in two comprehensive high schools in separate Michigan school districts. The case study design focuses on nine "focus teachers" equally distributed among three "focus PLCs" as embedded units of analysis; it is seeking to achieve an explanatory goal; and it provides for effective within and cross case analyses. Data collection and analysis occurred during the 2019-2020 school year as focus teachers, per job requirements, engaged with, and so responded to, the respective demands of the teacher evaluation and PLC structures of their district and school.

The purpose of this chapter lies in elaborating this dissertation's research design and the execution of its methodology, which have been briefly indicated. The chapter first describes the case study design and gives the rationale for using a case study method. The chapter's second section proceeds to the sampling strategy for the case high schools and follows to an identification of the individual research participants—focus teachers and administrators—together with the focus PLCs in the two districts/schools. The third and fourth sections explain, respectively, the data collection and data analysis processes that have been conducted. The final chapter section describes how the validity of the findings of this research are established.

Multiple Explanatory Case Study Design

Creswell (2013, p. 97 [author's emphasis]) describes a "case study" in terms of its *defining elements* that imply when a case study design is a good research fit:

> [A case study is a] qualitative approach in which the investigator explores a real-life, contemporary bounded system (a *case*) or multiple bounded systems (cases) over time, through detailed, in-depth data collection involving *multiple sources of information* (e.g., observations, interviews, audiovisual material, and documents and reports), and reports a *case description* and *case themes*.

(See also Patton, 2015; Miles, et al., 2014; Yin, 2014). Taking a somewhat different tact, Yin (2014, p. 9) specifies three *conditions* befitting of a case study design: "(a) the type of research question posed is a how or why question; (b) the researcher cannot manipulate the behavior of study participants; and (c) the study focuses on contemporary events" (see Baxter and Jack, 2008). Further, when Yin (2014) identifies three *types* of case study, he in effect distinguishes three possible *purposes* for a case study: (1) an exploratory case study advances research questions, hypotheses, and processes for future research to use or pursue; (2) a descriptive case study gives a detailed depiction of a phenomenon in its actual context; and (3) an explanatory case study presents a state of affairs, how or why it came to be.

An explanatory case study design readily fits the instant study. The research questions on their face not only anticipate Yin's three conditions, but also have answers with an explanatory purpose (2014, p. 9). The policies and policy context, moreover, can be studied in detail and in-depth as implemented by educators in real-life, multiple bounded systems called schools, and with the reporting of case descriptions and themes (Creswell, 2013; Miles, et al., 2014; see Baxter and Jack, 2008; Patton, 2015). The selected methodology, in sum, represents a strong design choice, as it satisfies the applicable criteria, however articulated.

This explanatory case study is a multiple or collective one because it gathers comprehensive data from each of at least two cases of interest for in-depth study and comparison (Patton, 2015; see Baxter and Jack, 2008; Stake, 2005; Yin, 2014). A primary advantage of this aspect of the research design is that "[m]ultiple-case sampling adds *confidence* to findings" (Miles, et al., 2014, p. 33; authors' emphasis). This advantage results in large measure from the opportunity to establish, compare, and confirm (or not) analytic findings from more than a single context or data set (Yin, 2014).

The key according to Baxter and Jack (2008), Yin (2014), and Miles, et al. (2014) is the replication of each case's results—here, how each case contributes to ascertaining and comparing teachers' agentic responses to PLC and evaluation demands as emerging from their engagement with, and as shaped by, interacting structures. In the Stake's terminology, this research falls on the "instrumental," as opposed to the "intrinsic," end of the continuum of multiple or collective case studies because the cases together are "examined mainly to provide insight into an issue or to redraw a generalization" (Stake, 2005, p. 445; see also Patton, 2015). Confidence in any generalization or comparison is further enhanced because the units of analysis are the focus teachers/focus PLC members embedded or nested in each case high school (Patton, 2015). As more fully explained in the following sections, the research design as proposed and as executed had a substantial probability of effectively enabling persuasive answers to the research questions.

Sampling Strategy and Participants

My sampling strategy consisted of a purposeful selection. As Maxwell (2013, p. 97; emphasis supplied) explains:

> In this strategy, particular settings, persons, or activities are selected deliberatively to provide information that is particularly relevant to your questions and goals, and that can't be gotten as well from other choices. … *Selecting those times, settings, and individuals that can provide you with the information that you need to answer your research questions* is the most important consideration in qualitative selection decisions.

My strategy was the type of purposeful selection that Patton (2015, p. 267) calls "comparison-focused" because it involved the selection of "cases to compare and contrast to learn about the factors that explain similarities and differences." A multiple case sampling design for

comparison and contrast should occur at the level of the phenomenon of interest (Miles, et al., 2014; Patton, 2015).

My research problem, questions and design together reflect that the phenomenon of interest is the interaction of teacher agentic responses to related PLC and evaluation demand structures as implemented in high schools confronting a significant coherence challenge. The research was to focus on whether and how teachers adapted their agentic responses, and on whether and how related demand structures shaped those responses *in the usual, ordinary course of a typical, "normal" school year.* That is, my research interest lay in certain basic *processes* that routinely play out in the context of teachers' regular work in regularly operating schools, and *not* in a context of initial implementation or of other special circumstances such as high accountability pressure or financial distress (see Lane, 2020). Indeed, the presence of unusual external forces in the case schools' context could undermine the goals of this research through an unrelated, but disproportionate, effect on the processes under study.

By the 2019-2020 school year, many districts and schools had institutionalized and routinized processes of teacher evaluation and PLC, along with even earlier reforms such as school level accountability systems. The schools and their districts still had ongoing coherence challenges, and teachers still had to navigate multiple demands. In the vast majority of these schools, however, teacher evaluation had proven to lack high stakes (see, e.g., Table 3.1; Kraft and Gilmore, 2017), and the professional accountability which PLCs were to develop was of a different order altogether (see, e.g., Adams and Kirst, 1999; Mϕller, 2008). Further, most of these schools posted mid-range results in terms of school-level accountability measures, being neither especially high nor low performers. Therefore, to keep the focus on the phenomenon of interest, and to obtain the data to answer the research questions, the case schools would be

intentionally drawn from this large set of "typical" schools, and so school contexts, in which teacher evaluation and PLC were routinely implemented each school year while not subject to disproportionate influence from any obvious external source.

I chose high schools as case schools for several reasons. First, high schools "occupy a strategic, intermediate position in the education system" (McLaughlin and Talbert, 2001, p. 125), especially in the era of stakeholder and policymaker demands for graduates' "college and career readiness." Yet, extensive reform attention to this school level is relatively recent. Second, due to such characteristics as size, organization, program, resources and composition, high schools present an unusually challenging (and interesting) context for the implementation of multiple reforms, including those involving accountability and professional community (e.g., Little, 2002b; Tichnor-Wagner, et al., 2016). In addition, the same distinctive characteristics offer more variation within each system—i.e., more PLCs over more content areas; more administrators evaluating more teachers' performances; more opportunities for the interaction of PLC and evaluation demands and responses. The degree of such variation may well heighten the coherence challenge or relating demand structures as it broadens the scope for teacher responses. Finally, my personal experience as a high school teacher would facilitate my understanding of and communication with focus teachers, my ability to build rapport with research participants, and my navigation of high school systems and processes.

The strategy of sampling similar cases but for the phenomenon of interest, which forms the basis of potentially revealing comparisons, constitutes an approach that Miles, et al. (2015, p. 32), following Goetz and LeCompte (1984), call "comparable case selection." It is a form of replication strategy commonly and appropriately employed in multiple case study research (see also Yin, 2014). For the reasons stated above, I wanted to apply the strategy while minimizing, to

the extent possible, significant differences in external factors, such as accountability pressure, resources, or teacher or student populations, that might distort or overly influence the functioning of the two teacher quality systems under study. The critical variable for attention, examination and comparison would instead lie in the "street level" functioning of and teachers' engagements with the related demands of PLC and evaluation. By design, this study could then be properly centered on the case-embedded units of analysis—namely, the focus PLCs with their respective focus teacher-members, whose interacting responses might play a role in addressing the coherence challenge. Moreover, this sort of sampling from a relatively large pool of potential case schools served practical considerations of participant recruitment and site access.

Accordingly, I sought a sample of two similar high schools, each with similarly formalized systems of grade level/content PLCs. As schools within traditional K- or pre-K-12 districts, the candidate comprehensive high schools would be of middling size and situated in similar middle-sized locales by area, population and density—i.e., neither urban nor rural. Both high schools would offer and provide the same or similar educational program and support services. Both would be at least broadly comparable in student and teacher demographic make-up, as well as in student achievement, teacher effectiveness, and school accountability rating measures. Thus, neither would be especially high performing or low performing; neither would be subject to high accountability pressure or threatened by accountability sanctions. Both would be adequately resourced with no experience or danger of financial distress. Of course, any two candidate schools would be implementing the state mandated evaluation system, requiring the same kind of engagement by their teacher-members, who would also be evaluated with respect to the same sort of subject matter curricula, pedagogical content knowledge and instructional performance, and student assessments and student growth measures.

With respect to PLCs, each case high school would have to be large enough to have grade-level/content PLCs with teacher membership determined at the beginning of the school year and remaining stable throughout it. Each PLC would have at least two members whose primary assignment was the same grade-level/content, although an overly large PLC, say with more than five members, would complicate comparisons and practicalities. The PLCs would have regular dedicated meeting times during the teacher work day, together with established expectations for participation, tasks and products, with some system of administrative supervision.

It would be helpful for within case comparison if the two grade-level/content PLCs in each high school were in different departments within the principal disciplines—English Language Arts (ELA), Mathematics, Science, Social Studies. It would be helpful for cross-case comparison to have PLCs from the same two departments in both high schools. These four departments tend to have larger and more stable teacher memberships. In addition, as state-tested subjects, state law specified the same student growth measure and weighting. These core subjects, particularly Math and ELA, were (and are) often figure most prominently in teacher quality, school improvement and accountability plans, measures and attention.

Based on the desired characteristics outlined above, I began several rounds of contacting by email and telephone contact districts' central office administrators (e.g., superintendents, assistant superintendents for curriculum and instruction, secondary education directors) and high school principals. I first concentrated on districts with high schools that (1) were within the appropriate range of size, (2) had an established system of PLCs organized by grade-level/content, (3) shared similar locale classification (e.g., suburban or town locale of similar size), and (4) were within a reasonable driving distance from my home or MSU. I identified

potentially appropriate districts, high schools and administrators through (a) personal knowledge from sixteen years of high school teaching and teacher leadership positions; (b) introductions and recommendations from contacts in K-12 education, especially among ISD educators and teacher union representatives; and (c) online research involving state-collected district/school data and district/school web sites. With each round of contacts, I expanded the range of potentially qualifying criteria listed above, except for (2), which often proved difficult to determine before initial contact and expression of interest.

Just after the student school year began in September, 2019, the administration in one district and high school agreed to access. With administrative support, the high school ELA department chair agreed to recruit her grade-level/content PLC (ELA 11). Of its five members, three (including the department chair) consented to full participation; the remaining two consented to observation in PLC meetings, but declined interviews and questionnaires for reasons of time and personal circumstance unrelated to the study. Because this PLC would be meeting in the third week of September, I immediately began working with the consenting teachers in the ELA 11 PLC. Meanwhile, I was under the impression that a Math PLC would similarly agree to participate. However, after a long delay in arranging communication and meeting with the department chair, and further delay in communications with the Math teachers, participation was declined. Requests to the Science and Social Studies departments where likewise unsuccessful. I thus continued with one PLC and three of its teacher-members as full study participants.

Eventually, in October, a second high school satisfying my criteria, particularly those pertaining to size and PLC model, agreed to provide access. The principal ensured that the

district consented, and he arranged me to meet with and secure consents from the three teacher-members of the ELA 10 and of the Math 10 (Geometry) grade-level/content PLCs.

As Table 4.1 shows, the two *districts*, Woodford Public Schools and Walker Public School District, are broadly comparable. (As noted in each table, all names are pseudonyms.) Although the first recruited, Woodford, is relatively larger and has a more racially/ethnically diverse student enrollment, other student demographics are similar in percentage terms. In addition, both report similar indicators of financial health and stability, including student enrollment numbers and a net positive benefit from being schools of choice. Their respective districtwide distributions of teacher effectiveness ratings both match the statewide figures (Table 3.1 above).

Table 4.1 Comparison of Participating District Data

Districtwide Data	Woodford	Walker	Data Year
State Region	South-Central	Southeast	n/a
Community Locale	Suburb	Suburb	n/a
Per Pupil Foundation Allowance	$8,200	$8,500	2019-20
Per Pupil Revenue	$11,400	$10,500	2018-19
Per Pupil Instructional Expenditure	$7,000	$6,200	2018-19
Total Student Enrollment	5400/5550	2910/2875	2019-20/2018-19
Male	50.0%	50.0%	2018-19
Female	50.0%	50.0%	2018-19
African-American	11.0%	6.5%	2018-19
Hispanic	14.0%	8.5%	2018-19
White	62.5%	83.%%	2018-19
Two or more races/ethnicities	11.%%	0.5%	2018-19
Economically Disadvantaged	44.5%	45.0%	2019-20
Students with Disabilities	12.5%	9.5%	2018-19
ELL%)	4.5%	2.5%	2018-19
Schools of Choice (Y/N)	Yes	Yes	2019-20
Resident Pupils Leaving	620	275	2019
Non-Res Pupils Attending	1330	1415	2019
Total Teacher Count	330	165	2018-2019
Teacher Rating Highly Effective or Effective	99.0%	98.0%	2018-19
Teacher Rating Minimally Effective or Ineffective	1.0%	2.0%	2018-19

Note: District names are pseudonyms, and some identifying data are rounded, omitted, combined, combined, or otherwise slightly altered from the source data to protect their anonymity.
Source: Michigan Department of Education, MI School Data, https://www.mischooldata.org

Table 4.2, a comparison of the participating *high schools* in terms of student demographics, shows that Woodford High School is larger and more racially/ethnically diverse than Walker High School. Walker, however, posts somewhat better measures of student achievement.

Table 4.2 Comparison of Participating High School Student Demographics

High School Student Data	Woodford HS	Walker HS
Total Student Enrollment	1700	900
Male	52.0 %	49.5%
Female	48.0	50.5%
Asian	3.5%	1.5%
African-American	11.0%	5.5%
Hispanic	13.0%	8.5%
White	62.5 %	83.0%
Two or More Race/Ethnicities	9.0%	0%
Economically Disadvantaged	35.0%	36.5%
Special Education	11.5%)	9.0%
English Language Learner	3.5%	1.5%
Proficiency on Grade 11 M-Step, all subjects	23.0%	35.0%
SAT Total Score	1000	1045
SAT College & Career Readiness Benchmarks	33.0%	41.0%
4-Yr Graduation Rate	87.0%	>95%
Drop-Out Rate	<5%	<5%

Note: All data is from the 2018-19 school year, except for M-Step proficiency percentage, which represents 2017-18 school year data; high school names are pseudonyms, and some identifying data are rounded, omitted, combined, combined, or otherwise slightly altered from the source data to protect their anonymity.
Source: Michigan Department of Education, MI School Data, https://www.mischooldata.org

With respect to the teaching staff demographics of the participating high schools, Table 4.3 reveals a high degree of similarity in percentage terms. Further, but for some variance in the ratio of highly effective to effective, the distribution of teacher evaluation ratings is practically the same.

Table 4.3 Participating High School Teaching Staff Demographics

High School Staff Data	Woodford HS	Walker HS
Total Teaching Staff	100	40
Male	41.0%	46.0%
Female	59.0%	54.0%
White	>99.0%	>91.0%
Tenured Teachers	70.0%	71.5%
Non-Tenured Teachers	30.0%	28.5%
Teachers with <1-5 years of experience	30	13
Teachers with 6-15 years of experience	25	12
Teachers with 16-25 years of experience	34	13
Teachers with 26->30 years of experience	11	2
Teacher Rating of Highly Effective or Effective	100%	100%
Teacher Rating of Minimally Effective or Ineffective	0%	0%

Note: All data from 2019-20 school year, except for teacher performance rating data, which represents 2018-19 data when Woodford HS had a teaching staff of 95, and Walker HS had a teaching staff of 47; high school names are pseudonyms, and some identifying school data are rounded, omitted, combined, or otherwise slightly altered from the source data to protect their anonymity.
Source: Michigan Department of Education, MI School Data, https://www.mischooldata.org

As explained above, I was able to secure participation by one grade-level/content focus PLC, ELA 11, in the Woodford district/high school, consisting of three fully participating focus teacher-members (as well as two observation only teacher-members). By contrast, in the Walker district/high school, I successfully recruited two grade-level/content focus PLCs—namely, ELA 10 and Math 10—each of which had three fully participating focus teachers-members. For the three total focus PLCs with nine total focus teacher-members, Table 4.4 provides comparative background information from their respective questionnaires:

Table 4.4 Participating Focus Teacher-PLC Member Information

Name	HS	PLC	R/ Eth	Gndr	Age	HiDgr	Tenure	Yrs Tch'g	Yrs Dist	Yrs HS
Ms. Dalmore*	Woodford	ELA 11	W	F	43	MA	Yes	21	21	21
Ms. Laphroaig	Woodford	ELA 11	W	F	38	BA	No	7	5	5
Ms. Macallen	Woodford	ELA 11	W	F	42	MA	Yes	20	20	20
Ms. Daniels	Walker	Math 10	W	F	26	BA	No	4	4	4
Mr. Bulliet	Walker	Math 10	W	M	24	BA	No	2	2	2
Mr. Roses	Walker	Math 10	W	M	39	BA	Yes	13	7	7
Ms. Trace	Walker	ELA 10	W	F	40	MA	Yes	15	15	15
Ms. Williams	Walker	ELA 10	W	F	46	BA	No	14	3	3
Ms. Forester	Walker	ELA 10	W	F	40	MA	Yes	17	17	17

Note: Teacher and high school names are pseudonyms.
*Ms. Dalmore is the ELA department chair, who has additional district and school teacher leader responsibilities, including with respect to departmental PLCs.
Source: Focus teacher questionnaires.

Finally, three administrators from each district were interview participants. Table 4.5 gives their names (for purposes of this study), position title, and years of experience in that position and in the district.

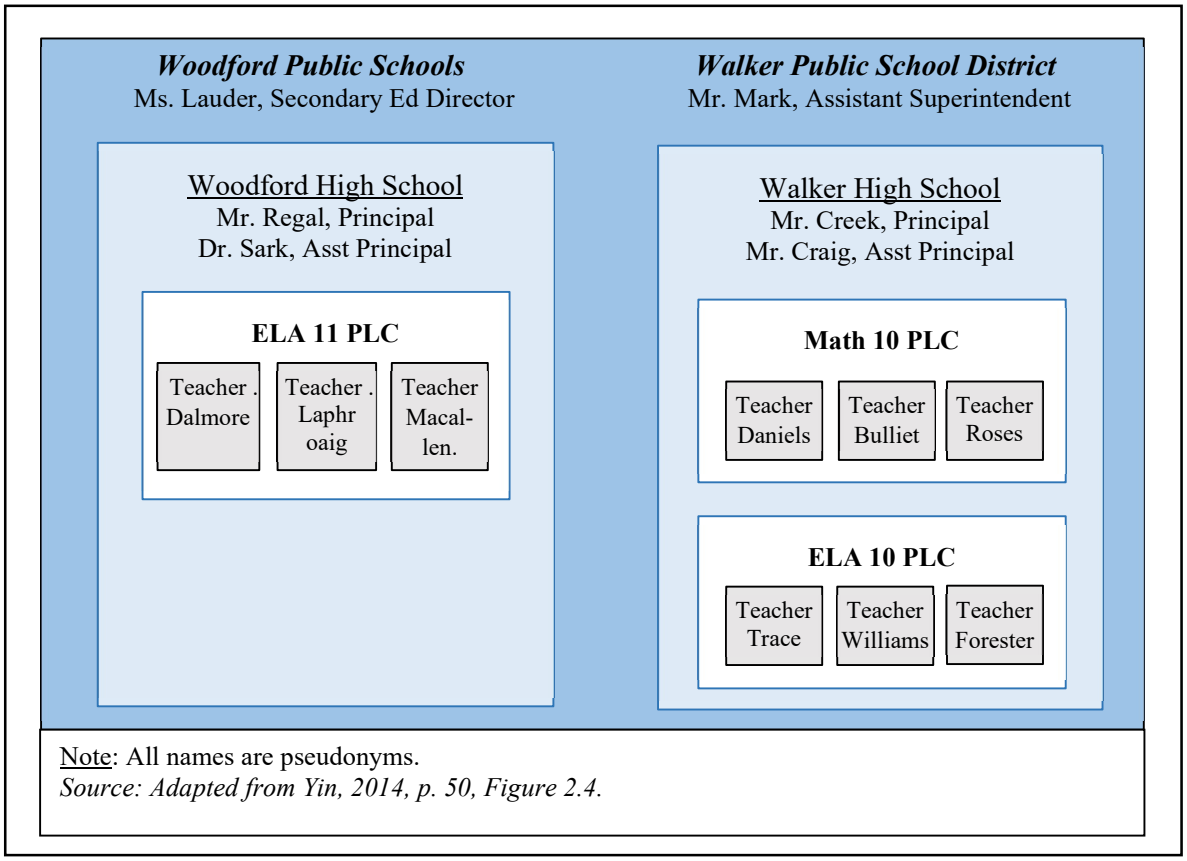
Table 4.5 Participating Administrator Information

District/HS	Administrator	Position	Yrs Position	Yrs District
Woodford	**Ms. Lauder**	Secondary Curriculum Director	4	17
Woodford	**Mr. Regal**	High School Principal	7	7
Woodford	**Dr. Sark**	High School Assistant Principal	3	3
Walker	**Mr. Mark**	Assistant Superintendent	2	8
Walker	**Mr. Creek**	High School Principal	7	20
Walker	**Mr. Craig**	High School Assistant Principal	4	23

Note: Administrator and high school names are pseudonyms.
Source: Participant interviews.

I summarize the organizational positions and relationships among the participants in Figure 4.1, which also reflects the multiple case study with embedded units design of this dissertation.

Figure 4.1 Embedded Multiple Case Study Units and Focus Teacher Participants



Data Collection

Baxter and Jack (2008, p. 554) call the use of multiple data sources "[a] hallmark of case study research," an accurate characterization for several reasons (see Creswell, 2013; Maxwell, 2013; Miles, et al., 2014; Patton, 2015; Yin, 2014). First of all, multiple sources are necessary to the development of an in-depth understanding of the cases requires multiple sources. Such diverse sources provide information about different aspects of the cases, helping to ensure that the phenomena of interest are studied through a variety of lenses or perspectives (e.g., Miles, et al., 2014). Second, the convergence of evidence derived from range of data collection forms and means strengthens case study understandings, explanations, and findings (Baxter and Jack, 2008; Maxwell, 2013; Miles, et al., 2014; Yin, 2014). Multiple sources and methods greatly improve the opportunities for identifying and developing what Erickson (1986) calls "key linkages"

among data sources, together with what have also been described as "propositions" (Yin, 2014) and "analytic generalizations" (Firestone, 1993). Third, the use of multiple sources not only increases the credibility of collected data (Baxter and Jack, 2008), but this strategy also is essential to triangulation that establishes the validity of research findings (Cresswell, 2013; Miles, et al., 2014; Yin, 2014).

The data on which case studies frequently rely are sourced in various forms of interviews and observations, together with documents, archival records and physical artifacts (Yin, 2014). Pursuant to my case study design, data sources for this dissertation were several: (1) focus teacher questionnaires, (2) videotaped and field-noted/memoed observations of focus PLC meetings, (3) recorded and noted/memoed in-person or phone interviews of focus teachers and participating administrators, and (4) contemporary document reviews. These types of data were the best fit to address my research questions, and the same types of data sources have been used by qualitative researchers investigating teacher responses to and engagement with formalized PLC and evaluation (e.g., Marsh, et al., 2017; Reinhorn, et al., 2017; Rigby, 2014; Valli and Buese, 2007). Although qualitative research scholars are not entirely consistent in the category labels for data collection sources, all agree that researchers commonly, appropriately and effectively rely upon the kinds of data I collected here (Baxter and Jack, 2008; Creswell, 2013; Erickson, 1986; Patton, 2015; Maxwell, 2013; Miles, et al., 2014; Yin, 2014).

Focus Teacher Questionnaires. Each teacher completed a questionnaire shortly after data collection began at each high school research site—i.e., late September, 2019 in Woodford and late October in Walker (copy of the questionnaire attached as Appendix A). Each part of the two-part questionnaire served a different purpose. The first part aimed to learn the focus teachers' overall views on the goals of, influences on, and the relationship between their

experiences of PLC and evaluation. I then used each teacher's answers, plus patterns in the answers among focus PLC members, to inform interview questions and observed expressions. In addition, by comparing questionnaire responses with interview answers and observed PLC work, the questionnaire played a significant role with other data sources in the data triangulation conducted during data analysis. The second part of the questionnaire collected background data on each focus teacher. Such data included basic demographic information, together with somewhat more specific information regarding each respondent's education, experience, and assignment.

Focus Teacher and Administrator Interviews. I twice interviewed each focus teacher at each case high school between October 15, 2019 and March 9, 2020. These interviews included all three fully participating teachers constituting the ELA 11 PLC at Woodford High School, and all six teachers, in two trios, constituting the Math 10 PLC and the ELA 10 PLC at Walker High School. All interviews were semi-structured in form, following one protocol for the first interview (Appendix B) and a separate protocol for the second interview (Appendix C), and all were conducted in one-on-one settings in person or by telephone. All interviews were audio recorded; in addition, I took notes for the duration of each interview, and I wrote memos with a more holistic, impressionistic assessment and main point summary of each interview as soon as practicable thereafter. The date, mode and duration of each interview for each focus teacher are given in Tables 4.6 and 4.7.

Table 4.6 Woodford Focus Teacher Interviews

Focus Teacher	PLC	Interview #1 Date	Mode	Duration (minutes)	Interview #2 Date	Mode	Duration (minutes)
Ms. Dalmore*	ELA 11	10.15.19	In person	57	01.28.20	In person	56
Ms. Laphroaig	ELA 11	10.09.19	Phone	46	01.29.20	Phone	37
Ms. Macallen	ELA 11	10.22.19	Phone	44	01.28.20	Phone	34

*Ms. Dalmore is the ELA department head, who has additional district and school teacher leader responsibilities, including with respect to departmental PLCs.

Table 4.7 Walker Focus Teacher Interviews

Name	PLC	Interview #1 Date	Mode	Duration (minutes)	Interview #2 Date	Mode	Duration (minutes)
Ms. Daniels	Math 10	12.04.19	In person	31	02.27.20	Phone	29
Mr. Bulliet	Math 10	12.12.19	Phone	39	02.18.20	Phone	27-30*
Mr. Roses	Math 10	12.19.19	Phone	50	03.09.20	Phone	29
Ms. Trace	ELA 10	12.12.19	Phone	40	02.25.20	Phone	24
Ms. Williams	ELA 10	12.18.19	In person	42	02.25.20	Phone	18
Ms. Forester	ELA 10	12.18.19	In person	31	02.25.20	Phone	15

* Due to Zoom user/researcher error, only the last 15.25 minutes audio recorded; there are researcher notes for the entire time.

Woodford ran a traditional two semester calendar for the school year. Thus, the interviews of the Woodford teachers occurred roughly around the middle of the first semester, with the second interviews occurring shortly after the beginning of the second semester. Walker, by contrast, used a trimester system, where the student composition of all classes changes each trimester, even as the curriculum of non-elective classes like Math 10 and ELA 10 proceed for the full school year. Walker teacher interviews took place within the first few weeks of the second trimester, and second interviews were completed withing the last few weeks of that trimester.

Despite the differing calendars, the objective in both schools was to conduct the initial interviews once the school year was sufficiently under way that routines had been established and both the regular evaluation and PLC systems were operating in the ordinary course. Thus, first interviews concerned focus teachers' recent and ongoing experiences with, perceptions of, and participation in evaluation and PLC (see Appendix B). First interviews also enquired into the influences on and of evaluation and PLC, including how the two related to each other and to teachers' engagement with both.

The objective of the second interviews was to allow enough time to pass since the first, as well as into the school year overall, such that teachers both could reflect on how evaluation and

PLC work had progressed (or not) through much of the year and would be able to anticipate the likely course of evaluation and PLC to the end of the year. Second interviews accordingly asked focus teachers to assess the year so far, including how their experiences, perceptions, participations around evaluation and PLC had developed or changed since the first interview and how the same compared to their expectations going into the year (see Appendix C). Second interviews further concerned focus teachers' awareness of or action regarding connections between evaluation and PLC, as well as their expectations for both during the balance of the year.

Finally, together with any clarifying question related to the first interviews, PLC observations, or documents, second interviews addressed a newly salient issue that had arisen, at Woodford concerning PLCs and at Walker concerning evaluation. In addition to sub-departmental grade-level/content PLCs, Woodford had also separately formed, tasked, and scheduled time for "collaboration time PLCs." These were organized by instructional domain (e.g., student feedback, formative assessment); teachers pick one to join, which results in much larger, cross-disciplinary groups. Because this kind of PLC came up in several first interviews (including administrator interviews), I asked about them in focus teachers' second interviews for purposes of comparison with the focus PLC. Walker's central administration notified principals in December before the holiday break that building administrators would for the first time be conducting "midyear review" meetings with their teaching staff. Mr. Creek and Walker High School duly notified his staff and created a written questionnaire for teacher completion as the basis for these individualized meetings. Focus teachers' second interviews therefore included questions about this seeming change in the individual teacher performance evaluation structure.

I conducted one interview of three administrators in each district. In both Woodford and Walker, one interviewee was a central administrator with evaluation and PLC responsibilities, one was the high school principal, and one was an assistant high school principal. Like the focus teacher interviews, all administrator interviews were semi-structured following their own protocol (Appendices D and E). All occurred in one-on-one settings, and all were audio-recorded, whether conducted in person or by phone. I took notes through the full duration of each interview, and I wrote a memo soon after each ended. Tables 4.8 and 4.9 present the relevant details of these interviews.

Table 4.8 Woodford Administrator Interviews

Name	Position	Interview Date	Mode	Duration (minutes)
Ms. Lauder	Secondary Director	10.15.19	In person	29
Mr. Regal	HS Principal	11.27.19	Phone	37
Dr. Sark	HS Asst Principal	01.16.20	Phone	34

Table 4.9 Walker Administrator Interviews

Name	Position	Interview Date	Mode	Duration (minutes)
Mr. Mark	Asst Supt	01.22.20	In person	44
Mr. Creek	HS Principal	12.11.19	In person	43
Mr. Craig	HS Asst Principal	12.11.19	In person	46

The duties of the participating central administrator in Woodford, who was a high school teacher and instructional coach before assuming her present position, included secondary curricula, instruction and assessment, as well as supervisory authority over teacher evaluation and professional development, which included PLCs. The participating assistant superintendent in Walker had a similar role with respect to teacher evaluation, and he worked with building principals on PLCs; he also was responsible for district data collection and provision, including that necessary to evaluation and professional development. Thus, the relevant interview protocol (Appendix E) mainly focused on the structures of evaluation and PLC, including how the district

had adopted and implemented evaluation and PLC, how both worked and the challenges each presented, and the leadership and support provided by district. These interviews also addressed any connections or interactions between evaluation and PLC, as well as the administrators' perceptions of the effects of both in the district's high school.

The principal and assistant principal were the only building administrators at Walker High School, so they jointly implemented the structures of PLC and evaluation. At Woodford, building administration included the principal and four assistant principals. The principal and the assistant principal who was interviewed, Dr. Sark, together conducted the majority of teacher evaluations; both were also involved in monitoring and supervising the PLCs, with Dr. Sark serving as the administrative liaison to (among others) the ELA department and its sub-department grade-level/content PLCs. At both high schools, the purpose of the building administrator interviews (Appendix D) lay in understanding how evaluation and PLC were implemented at the school level. For example, interviewees were asked about how the two systems worked in practice, what the challenges of implementation were, and how teachers seemed to perceive evaluation and PLC. Additionally, building administrators were asked to identify and explain any connections between PLCs and evaluation, and to compare and contrast the effects of each in their high school.

Observations of PLC Meetings. The opportunities for PLC observations differed within and between the two case high schools. Both Woodford and Walker organize their PLCs, and schedule PLC meetings, by disciplinary department. However, the principal departments at both schools (ELA, Math, Science, Social Studies) have sub-departmental PLCs based on grade-level/content, the embedded units of study here. Both schools regularly schedule departmental

meetings; but because departmental tasks or activities must be completed first, there is no fixed or consistent amount of time for grade-level/content PLC to work.

Further variance in observational opportunities arise from how each school schedules PLC time. At the secondary level, Woodford had a late start (no students) period practically every Wednesday during the semester. However, the district calendar allotted this time to several kinds of professional learning—whole building initiatives, cross-disciplinary "collaborative time" teacher teams (noted above), "content time" (departmental) PLCs, cross-school teacher meetings. As a result, the sub-departmental grade-level/content PLCs only met about four times a semester. In Walker, time for departmental/PLC meetings was obtained by running a modified schedule every Wednesday, although departments only met every other Wednesday, since a Response to Intervention (RTI) program occupies the time during alternate Wednesdays. Walker had also converted some after-school staff meetings to departmental/PLC time.

It was in these contexts that, between late September 2019 and late January 2020, I observed the Woodford ELA 11 PLC on four occasions for a total of slightly more than 160 minutes. At Walker High School between late October, 2019 and mid-February, 2020, I observed the ELA 10 PLC three times totaling about 200 minutes and the Math 10 PLC five times for a total of nearly 160 minutes. For all PLC observations, I sat apart from the group that was meeting while videotaping the meeting using my smart phone with a specialized application interfacing with a flash drive for video file recording and storage. In addition to the video recording, I made field notes for the full duration of each observation. After each observation, I wrote a memo which included more holistic impressions and reflections, as well as salient points for later consideration or follow-up.

Table 4.10 details the observed meetings of the Woodford ELA 11 PLC.

Figure 4.10 Woodford ELA 11 PLC Observations

Date	Present	Absent	Duration (minutes)
09.25.19	Dalmore, Laphroaig, 2 interns**	Macallen, NFPT A/B* (B w/ELA 10)	64
10.30.19	Dalmore, Laphroaig, Macallen, 2 interns	NFPT A/B (B w/ELA 10)	32***
01.08.20	Dalmore, Laphroaig, Macallen, NFPT A/B, 2 interns	None	28
01.22.19	Dalmore, Laphroaig, Macallen, NFPT A, 2 interns	NFPT B	38

* NFPT A/B = there were two "not fully participating teachers," who were members of the PLC but consented only to observations, not interviews or questionnaires.

**Pre-service interns = two of the focus teachers, Dalmore and Macallen, and one of the NFPTs had pre-service interns who sometimes attended the PLC meetings; with the approval of their mentor teachers; all three consented to observations.

***Approximately 6 of the last 8 minutes of this meeting not video recording due to video mal-function; field notes were taken for entire meeting time.

It is significant that, of the four PLC meetings scheduled for the first semester, the last three were observed (the first occurred before students started), together with the first meeting of the second semester. Furthermore, with one exception, the three focus teacher-members attended the meetings for their full durations. "Not Fully Participating Teacher A" (NFPT A), who I understand had health issues, only attended two of four meetings, and NFPT B only one of four, in part because her primary PLC was the ELA 10 grade-level/content group. At all meetings the three focus teachers were the most active participants. Ms. Macallen's and NFPT A's interns attended all meetings, with the latter especially participating in the discussions. Finally of note for data collection purposes, there was one video glitch during the third meeting.

Tables 4.11 and 4.12 provide information like Table 4.10 with respect to the Walker PLCs, ELA 10 and Math 10.

Table 4.11 Walker ELA 10 PLC Observations

Date	Present	Absent	Duration (minutes)
10.30.19	Trace, Williams	Forester**	n/a*
11.20.19	Trace, Williams, Forester	None***	84^
01.22.20	Trace, Williams	Forester	69^
02.19.20	Trace, Williams	Forester	46^

*All PLC time used for departmental business; no grade-level/content PLC meeting time.

**Forester's primary PLC based upon her assignment was in the science department.
***Forester was present for about 60 of the 84 minutes of this meeting.
^Meeting after school voluntarily extended beyond scheduled PLC time period.

Table 4.12 Walker Math 10 PLC Observations

Date	Present	Absent	Durations (minutes)
11.13.19	Bulliet, Roses	Daniels	21
12.04.19	Bulliet, Roses, Daniels	None	n/a*
12.11.19	Bulliet, Roses, Daniels	None	n/a*
01.16.20	Bulliet, Roses, Daniels	None	25***
01.20.20	Bulliet, Roses, Daniels	None	75**
02.05.20	Bulliet, Roses, Daniels	None	11
02.06.20	Bulliet, Roses, Daniels	None	25***

*All PLC time used for departmental business; no grade-level/content PLC meeting time.
** This PLC meeting occurred as a part of a full day of professional development. Although the meeting was videoed, a video application malfunction at the end resulted in the loss file; field notes were taken for entire meeting time.
***Voluntary Math 10 PLC meetings during focus teacher-members' common lunch period.

Note that, on the one hand, both of the Walker grade-level/content PLCs "lost" meeting time when their respective departmental meetings took all or most of the entire available meeting time. On the other hand, both made up time through voluntary meetings during a common lunch period (Math 10) or after the regular school day (ELA 10). Finally, both of these PLCs had a maximum of three focus teachers. Because Ms. Forester was assigned more Science than ELA classes, she usually went to the Science Departmental/Biology PLC meetings. As a consequence, she attended only one of the observed ELA 10 PLC meetings. The Math 10 PLC meetings were the reverse. Although she attended the Algebra PLC meeting at the end of the November 13, 2019 Math Department time, thereafter Ms. Daniels and the other Math10 focus teacher-members attended all observed meetings. During the longest of those, an unexplained technical failure in video application software resulted in the loss of the entire video file of that meeting.

Document Reviews. I reviewed the publicly available information about the districts and highs schools that was available online at their websites and though the state department of education. This provided extensive background data on district/school governance, finance,

students, staffing, accountability measures, and other basic information. More specifically, online postings identified and gave access to documentation of and documents used in the evaluation system in each district/school, regarding, for example, the Marzano evaluation tool used in Woodford and the Danielson tool used in Walker. For the 2019-20 school year, Walker published and distributed an "Evaluation Handbook" to all teachers and administrators, which included documents explaining and forms to be used in teacher evaluation. Supplemental and explanatory information about the Marzano Focused Teacher Evaluation Model adopted in Woodford and the Danielson Framework for Teaching adopted in Walker were also available from the state MDE or the district. Each focus PLC and several focus teachers individually provided access to documents showing its work in recently past and the current school year. These documents comprised PLC products such as (1) common curriculum standards and related common assessments, (2) rubrics and student learning progressions or objectives, (3) "course packs" and descriptions, and (4) lesson planning and materials. Particular PLCs made available documentary evidence regarding meeting formats, agendas, task and activities.

Data Analysis

Well-conducted, effective data collection and ongoing analysis leading to robust, highly trustworthy findings starts with a "researcher-as-instrument" who possesses characteristics like a solid familiarity with their study's phenomenon and setting of interest, and good investigative skills applied with an attention to detail and appropriate dispositions toward participants (Miles, et al., 2015, p. 42). Indeed, Yin directly addresses one of his principles of high-quality analysis to the researcher: "you should use your *prior, expert knowledge* in your case study" (2014, p. 168; emphasis in original). Before entering my doctoral program, I taught high school social studies for 16 years in a large K-12 school district in southeastern Michigan. During my tenure, I also

served in a number of teacher leadership positions and as an instructional coach. In the mid-2000s, I was part of the high school group that went through DuFour PLC training and implemented that PLC model first in our school; later it was adopted throughout the district. I was a teacher-member of Michigan Teacher Tenure Commission when the state began its teacher evaluation reforms, and I was subject to the state-mandated system for the last two years of my employment. While in my doctoral program, my research coursework, research assistant assignments, and practicum all included substantial elements of and experience with qualitative research.

Beyond prior knowledge and experience, the researcher conducting qualitative analysis proper should first aim to thoroughly know their data by reviewing it repeatedly, while also writing annotations, notes and memos (Bazeley, 2013; Creswell, 2013; Maxwell, 2013). Maxwell emphasizes that although memos serve many functions, they are "an essential technique for qualitative analysis [references omitted]" (2013, p. 105; see Miles, et al., 2014). As discussed in the data collection section above, I collected and wrote about the early data in the focus teacher questionnaires. Although interview and observation data were recorded, I made fieldnotes during each, and after each interview and observation I wrote a summary memo. I recorded the memos in a research journal, where I also wrote reflections, notes and "jottings" (Miles, et al., 2015), as well as, later, other kinds of process and analytical memos (Bazeley, 2013; Silver and Lewins, 2014). Further, after each interview and observation, I reviewed the audio/video recording, any documents obtained at the time, and my accompanying field notes, the last of which I annotated based on the first two.

As I completed a set of observations for each focus PLC and/or received relevant documents from each PLC's focus teachers, I re-watched their observational videos and

reviewed the documents. I took notes and made memos on the content and potential significance of both, among other research journal reflections on them. The video review process included noting the time range and content of potential excerpts for illustrative, corroborative and triangulation purposes. I marked and annotated copies of the documents for the same purposes.

I used Trint, a commercial online automated transcription service, to make a first draft of each interview transcript, which I reviewed for likely errors in content and formatting. I then again listened to the audio recording of each interview as I corrected each transcript. I reviewed and annotated each corrected transcript, and I also memoed and jotted in my research journal.

During the later rounds of review of all forms of collected data, I began to apply what Bernard, Wutich, and Ryan (2017, pp. 162-163) call a "constant validity check"—i.e., with intention, watching for disagreements among participants and checking participants' accuracy, attending to negative evidence and looking for alternative explanations, and examining one's theory to try and fit disconfirming evidence in (2017, pp. 162-163; see Miles, et al., 2014). In addition to negative or disconfirming evidence, I attended to what was *not* said and who did *not* participate (Guest, MacQueen, and Namely, 2014).

I coded and analyzed all interview transcripts using Dedoose (Salmona, Lieber, and Kaczynski, 2020; see Silver and Lewins, 2014). I first prepared and imported district, high school, PLC, and teacher descriptor sets, along with the fields each encompassed. I uploaded as well the interview transcripts for coded excerpting and the observational videos for illustrative excerpting. Next, I began my first cycle descriptive coding (Miles, et al., 2015) using a priori "structural codes" (Guest, et al., 2014), a.k.a. "organizational categories" or "topics" (Maxwell, 2013), which were developed primarily from the interview protocols and surveys. These initial codes also constituted the first iteration of my codebook. One round consisted of coding for

"evaluation," "PLC," and "relationships/inter-relationships"; the second, "teacher participation" and "administrative implementation." These codes represented the cross-cutting distinctions between and identifications of evaluation versus PLC structures and teacher responses, plus relationships and interactions (or their absence) among structures and responses. During the first cycle process, In Dedoose, I wrote memos and tagged data, "jotted," and noted potential child codes to reflect emerging themes within each topic and patterns across several topics (Bazeley, 2013; Silver and Lewins, 2014; Salmona, et al., 2020).

Referencing Stake (2006), Patton (2015, p. 536) points out that, in contemporary research, qualitative analysis is significantly directed toward "[a]nalyzing patterns and identifying themes across multiple case studies" (see Miles, et al., 2014). I accordingly proceeded with second cycle or pattern coding (Bazeley, 2013; Miles, et al., 2014; Patton, 2015). Having initially established initial parent codes categorizing the interview data, I recoded each code's group of excerpts to address any discrepancies between parent codes and to identify emerging themes and thematic patterns from which to develop sub-categories or child codes (Maxwell, 2013). I was at the same time expanding my codebook while refining particular code descriptions and relationships. For example, under the parent codes of evaluation structure-implementation and PLC structure-implementation, it became apparent that the implementation of both varied by the roles played by central administration versus building administration. Hence I added two child codes—"district administrative implementation," "school administrative implementation"—under each parent. In a similar way, the excerpts under the parent code of "PLC participation" supported distinct themes/child codes such as "PLC participation purpose" and "PLC participation value." I conducted a final round of coding, not only to review the prospective child codes for specificity and consistency, but to surface any significant facets of

them that as sub-themes of grandchild codes represent more focused points, often in participants' own wording (Miles, et al., 2014). For example, the child code "PLC participation work" encompassed data distinguishing between administration-driven and teacher-driven work.

I used the Dedoose "Chart Selector" functionality (Salmona, et al., 2020) to conduct horizontal and vertical code frequency analyses. I also conducted a coverage analysis by comparing sets of coded excerpts (Silver and Lewins, 2014). All frequency analyses were applied to and across groupings of individual participants, PLCs and high schools. I examined and compared code frequency totals, distributions, and patterns—whether expected or not, whether expected but absent—to assess the scope, significance, and utility of each child and related grandchild codes. In addition, I sought a high degree of fit or congruence among applied codes, emergent themes, and the supporting data by reviewing and comparing code descriptions together with and the coverage and content of coded excerpt sets. As a result of these analyses, I flattened and simplified the coding scheme and resolved any discrepancies when I merged several grandchild codes, and eliminated the "PLC participation" and "evaluation participation" parent codes (Bazeley, 2013; Silver and Lewins, 2014). I then revised the remaining code descriptions and recoded affected excerpts. The upshot was a codebook comprised of eleven parent codes—five PLC parents, five evaluation parents and one parent for relationships/inter-relationships. Each parent code had only two or three child codes (see Appendix F).

To assess and improve the reliability of my coding scheme and process, I used the inter-rater reliability (IRR) "Training Center" in Dedoose (Samona, et al., 2020). I created two tests, one for the parent evaluation codes and one for the parent PLC codes; both tests included the most important parent code—PLC/Evaluation Relationships and Inter-Relationships (Appendix F). The evaluation code test had 53 excerpts, the PLC code test 51; in both tests the excerpts

were not selected in any particular order, but were distributed approximately equally across the relevant codes and all interview transcripts. I did not review any of the entire excerpts when I selected them for the tests. The total of 104 excerpts represented slightly more than 10% of the total number of all excerpts (1006). I would take each test once, analyze the results to implement corrections and improvements, and then take each a second time (see Baxter and Jack, 2008).

I aimed first to experience how the testing functionality worked, mitigate the chance of memory bias, and gauge the intuitive precision of the codes, revealing the more patent code discrepancies. I thus piloted the evaluation and PLC tests "cold"—i.e., without any preparation, including any review of any excerpt source or larger context, or any review of the code definitions—on separate days after the creation of each. The resulting pooled Cohen's Kappa statistics were in the "good agreement" (evaluation parent codes) and "fair agreement" (PLC parent codes) ranges. The overall result in each test was most adversely affected by a significantly lower score with respect to one code (*not* relationships/inter-relationships).

I began analyzing the discrepancies at the lowest sub-score—i.e., with the code with the largest number of "disagreements" and so the lowest Cohen's Kappa statistic. I reviewed each excerpt in context of the broader interview and the identity of the interviewee (neither of which the test when taken shows). With such addition information, I then compared the excerpt to the two "disagreeing" codes' definitions. I discovered several silly mistakes resulting from how I first took the tests, as well as two excerpts that should not have been included in the tests to begin with—namely, a double-coded excerpt and an un-coded excerpt. More substantively, some disagreements could be resolved given the excerpt's larger context, others through refining the code description, and still others by both. Some resolutions required recoding; and whenever I modified a code description, I rechecked all the excerpts with the described code, recoding if

necessary. I iterated the process to examine and resolve all coding disagreements and to review and clarify the related excerpts, excerpt contexts, and coding descriptions.

In light of what I had learned about the testing process, and given the remedial steps taken, I retook each test after waiting a number of hours (the next day in one case). The resulting pooled Cohen's Kappa for the evaluation parent codes test was 0.98, and for the PLC parent codes test was 0.95, both being within the “excellent agreement” range. (Each test statistic may somewhat overstate overall agreement due to the inadvertent errors in test construction mentioned above.) In addition, the Cohen’s Kappa for each code in the first test was in one instance at most 0.07 of the others; and in the second test, the maximum between any two codes was 0.09, again in only one instance. The codebook and coded excerpts were ready for the application of more formalized analytical processing.

I employed the analytical tools in Dedoose in two inter-related ways (generally, Samona, et al., 2020). I used them, first, to identify and analyze themes and patterns in the data, as well as existing or missing relationships among the data (Bazeley, 2013; Maxwell, 2013; Miles, et al., 2014). These uses included an array of within and between case comparisons (Bazeley, 2013; Miles, et al., 2014; Patton, 2015; Yin, 2014). Second, Dedoose enabled me, in effect, to test tentative, alternative thematic interpretations, inferences, and propositions against what the data showed (Bazeley, 2013; Silver and Lewins, 2014; see Erickson, 1986; Maxwell, 2013). This testing process included a version of the analytic technique Yin describes as “explanation building” (2014, pp. 147-150; see Bazeley, 2013; Miles, et al., 2014). I applied, for example, filters for unit level (high school, PLC, individual participants) and codes to create sets of “Descriptor Fields x Codes Grid” charts to compare, say, the focus PLCs by predominance of district versus school PLC implementation, and administration versus teacher determined PLC

work. I would then make a set of such charts by individual focus teacher to see if a particular focus PLC member inordinately affected the PLC results. Again, at the PLC and focus teacher levels and filtering codes, I made several sets of "Code Frequency Descriptor Bubble Plot" charts to see, for instance, how relationship/inter-relationship codes varied among focus teachers by the relationship between school and district evaluation implementation. As a final example, I also ran a range of "Code Co-Occurrence" charts, which was telling for what they showed—and did not show—about the relationships by PLC, administrator, and focus teacher among the relationships/inter-relationships child codes and the PLC or evaluation purpose and value codes (Samona, et al., 2020; Silver and Lewins, 2014).

Analytical techniques "represent ways of *linking data to propositions*[.]" (Yin, 2014, p. 36; emphasis in original). In addition to the analytical moves summarized above, I further developed my thematic propositions, while addressing disconfirming evidence and potential alternatives, through several other steps calculated to establish and weigh the "key linkages" by which the data in its various forms and contexts supported the emerging assertions (Erickson, 1986; see Bazeley, 2013). These steps at the same time constituted a process of triangulation and corroboration among all data sources (Crestwell, 2013; Miles, et al., 2014). To these ends, I exported from Dedoose and printed sets of coded excerpts, which I reviewed, marked and annotated. I reviewed my Dedoose and research journal memos and other writings, distilling their ideas, insights and analysis into a consolidated set of notes for review, marking and annotating. I then compared the excerpts and notes to the other data—documents, questionnaires, and observations, together with, when appropriate, the interview transcripts.

At the beginning of this section, I quoted one of Yin's indicators of high-quality analysis (2014, p. 168), which this study satisfies. I have likewise sought to implement Yin's other three

principles as I conducted my analysis—i.e., I endeavored to attend to all the evidence, deal with all plausible rival interpretations as far as possible, and address the most significant aspects of this study (2014, p. 168). The reporting of my findings, with explanation and evidence, in the chapters that follow reflect these efforts, which at the same time contribute to the validity of my results.

Establishing Validity

According the Maxwell, "validity" in the context of qualitative research refers to "the correctness or credibility of a description, conclusion, explanation, interpretation, or other sort of account" (2013, p. 122). The overall validity of this dissertation can be established through the application of several layers of commonly accepted strategies to ensure the trustworthiness and authenticity of its research processes and findings (Baxter and Jack, 2008; Creswell, 2013; Glesne, 2016; Maxwell, 2013; Miles, et al., 2014; Yin, 2014). Rooted in key elements of the research design, these strategies undergird and reinforce study validity with high-quality implementation during participant recruitment, and data collection, analysis and reporting (e.g., Bazeley, 2013; Maxwell, 2013; Miles, et al., 2014). In this study, these elements ranged from the fit among the research questions, conceptual framework, and case study design, to data collection from multiple cases and involving multiple kinds and sources of data, to the well-implemented means and procedures for data analysis. All such elements, described in detail above, were integrated to ensure the overall quality of this research and its results (Baxter and Jack, 2008).

The design and execution of this study likewise addresses all three commonly distinguished types of validity—internal validity, external validity, and reliability (Miles, et al., 2014; Yin, 2014). First, the *internal validity* of a case study largely concerns its inferences and explanations for the phenomenon of interest (Yin, 2014). As Maxwell (2014) points out, a

number of scholars have proposed checklists of internal validity strategies, not for verifying conclusions, but for testing them, together with threats to them (e.g., Miles, et al., 2014: Patton, 2015). This study satisfies criteria Maxwell (2014, pp. 125-29) posits because it involved:

- collecting *rich data* through, for example, multiple, electronically and manually recorded interviews and observations of an array of educators in their work contexts over months of the school year (see also Yin, 2014);
- *searching for discrepant evidence and negative cases*, by, for example, repeatedly reviewing and writing about all data, and by systematically coding and comparing collected evidence (see also Cresswell, 2013; Patton, 2015);
- engaging in *triangulation* over a broad range of data sources and collection methods (see also Cresswell, 2013; Miles, et al., 2014; Patton, 2015)
- using *numbers*, derived from code frequency analyses, and to ascertain, aggregate and test the amount and variety of evidence supporting linkages or not (see also Guest, et al., 2014); and,
- making explicit *comparisons*, among, for example, individual and organizational participants at different levels, as well as making within and between case/embedded unit comparisons (see also Miles, et al., 2014; Patton, 2015).

In addition, the findings of this study are supported the appropriateness and quality of its analytical techniques, such as explanation building and cross-case synthesis (Yin, 2014).

In contrast to internal validity, *external validity* concerns the extent to which conclusions from particular research are generalizable beyond the immediate case study (Yin, 2014; see Miles, et al., 2014). Again, the present study's design and conduct included a number of methodological elements that enhance its conclusions generalizability. It is, for example, a

multiple case study, the purposeful sampling strategy for which contemplated a form of replication logic that still allowed for adequate comparison between cases and among focus PLCs and focus teachers (Baxter and Jack, 2008; Firestone, 1993; Miles, et al., 2014). Further, the prospect for external validity is similarly improved by its "how"-based research questions seeking to understand potentially transferable processes (Maxwell, 2013; Yin 2014). This follows because case study generalizability often rests on "the development of a theory of the *processes* operating in the case studied, processes that may well operate in other cases, but that may produce different outcomes in different circumstances [references omitted]" (Maxwell, 2013, p. 138; emphasis in original; see Firestone, 1993). The instant multiple case study seeks answers to research questions predicated on *how* teachers adapt their responses to the related structural demands of evaluation and PLC and on *how* those related structures shape those adaptive responses. This research thus required the investigation of interacting processes: teacher evaluation and PLC as implemented and teacher engagements in both.

Finally, the question of *reliability* asks whether a future researcher conducting the same case study by following the original design and methodology would likely arrive at the same conclusions (Yin, 2014; see Miles, et al., 2014). Through several means, this study has important indicia boosting its reliability. There was (and is) a congruence among the research questions, design, and data collection and analysis, with appropriate intercoder agreement checks made with at least adequate results; data was collected over a complete array of relevant settings, times and participants, and multiple participants' accounts converge when one would expect (Miles, et al., 2014; see Creswell, 2013). Data collection and analysis protocols were adhered to as well (Yin, 2014).

Whether considered holistically or as a sum of its component conceptual types, the validity of this study rests on multiple supports appropriate to its objectives, methodology, and findings.

CHAPTER 5
TEACHERS USE AGENTIC STRATEGIES TO ADAPT RESPONSES TO THE RELATED STRUCTURAL DEMANDS OF PLC AND EVALUATION

This chapter and the next set forth the research findings of this dissertation, all of which are guided by its structure-agency framework and research questions. This chapter concerns the agency branch of the framework, as its findings answer the first research question: How, if at all, do teachers adapt their agentic responses to performance evaluation and professional learning community demands as they engage with the structures of both?

The initial section below approaches the answer by describing the respective demands of evaluation and PLC at each school and in each district. I then identify the relationship found between teacher and PLC responses to those demands in the course of a school year as one of "punctuated compartmentalization." I describe both the "what" of compartmentalization as teachers normally segregate responses and the "when" of opportunities for punctuation occur as teachers under certain circumstances may at interact responses. The second section shows *how* teachers punctuate the usual state of compartmentalization, how teachers take opportunities to adaptively interact their responses. Within and across the cases, I find that teachers enact punctuation when evaluation and PLC demands interact by drawing on several adaptive response strategies. This chapter's final section presents findings about those teacher motivations and goals that result in the observed relationship of teacher responses. This section concerns the "why" of punctuated compartmentalization, which turns on teachers' assessments of the purpose and value of their engagement with PLC vis a vis evaluation.

"Punctuated Compartmentalization" of Teacher Responses to PLC and Evaluation Demands

Woodford and Walker teachers must respond, first, to the structural demands of PLC as implemented by their respective administrations over the course of each school year. As Chapter 6 discusses in depth, Woodford and Walker differ markedly in their PLC design and goal

orientations, which in turn shape the nature of teacher agency in response. Nevertheless, at both case sites, teacher agentic responses to PLC demands were manifested by how they participated in and contributed to their PLC's collective work, including the completion of PLC tasks and creation of PLC products.

For the 2019-2020 school year, as in the previous two, Woodford high school teachers belonged to two distinct kinds of PLCs, known as collaborative time and content time PLCs. The first kind were organized and tasked by instructional domain (e.g., student feedback, formative assessment). These corresponded to the Marzano instructional framework or model, which the district had adopted as the basis for its district-wide professional growth plan and teaching practice expectations. The district at the same time had selected a version of this framework, the Marzano Focused Teacher Evaluation Model, as its state-approved evaluation tool. Teachers picked one collaborative time PLC to join for the school year, which by design resulted in much larger, cross-grade, cross-disciplinary groups. The second kind of PLC was organized by disciplinary department, but they were subdivided into grade-level/content PLCs, such as the ELA 11 PLC, the Woodford PLC here under study. These PLCs concerned themselves with assigned matters involving disciplinary/grade-level curriculum, pedagogy, and assessment. At the beginning of each school year, the district set the schedule for all required PLC meetings and assigned each PLC certain work goals, tasks and products. The district-established calendar and assignments were implemented through building administration and department chairs. Central administrators periodically monitored PLC products, on which they gave feedback, which was used to prepare for the following year's PLC work.

Pursuant to the DuFour PLC model, which Walker high school had implemented as its own initiative several years prior to 2019-2020, teachers were organized into grade-level/content

PLCs, like the Math 10 and ELA 10 PLCs participating in this study. However, each regularly scheduled PLC time began with teachers meeting by department, where they engaged in required administrative tasks and regular DuFour video training before breaking into their grade-level/content groups. The high school ran its own shortened class schedule each Wednesday, with the "extra" period during the student day alternately used for required departmental PLC meetings and Response to Intervention (RTI) student meeting time (another high school initiative). In addition, the high school principal created after-school meeting time by converting some monthly staff meetings to PLC time. Following the DuFour model, building administration largely allowed PLCs to develop and work in their own way and at their own pace, provided they met certain broad expectations for collective activities and work products. In 2019-2020, the district for the first time began to assign particular tasks to departmental PLCs.

The Woodford and Walker PLCs responded to district and school structural PLC demands. They regularly met to continue work on the same sort of collective tasks and products over the school year. In contrast, evaluation demanded individualized teacher responses to state-mandated, district- and school-implemented structures. As outlined in the policy context section of Chapter 3 above, the evaluation system requires that teachers (and administrators) perform disparate tasks generating a variety of products with idiosyncratic timelines varying with each school year. In both Woodford and Walker, evaluation demands constrained teacher agentic responses relatively more than PLC structures. Yet teachers could exercise agency in how they engaged with evaluation components, including the timing of their formal observations and collection of student growth data, how they prepared for and enacted their roles in meetings and when observed, and in their selection among lessons and assessments, as well as their creation of other artifacts. The very nature of PLC and evaluation as implemented at both case schools

meant that focus teachers' agentic engagement with each *could* connect or otherwise relate responses to PLC and evaluation demands in some way and to some degree. I find that focus teachers actually *did* so, for the data reveal a strong, consistent relationship between focus teachers' responses over the course of a school year.

I call the overall relationship between focus teachers' agentic responses to evaluation and PLC demands "punctuated compartmentalization." As the phrase suggests, it consists in two alternative states. In the ordinary, day-to-day course, the more collective, regular and consistent demands of PLC were not usually relevant to the individualized, intermittent, and variable demands of evaluation. Teachers therefore kept their responses separate and independent. During the compartmentalized state, teachers addressed each immediate demand stimulus by directing a self-contained response only to whichever kind of demand confronted them at the particular point in time. A PLC revised a common assessment or lesson activity, for example, when a regularly scheduled PLC meeting took place. When a teacher received an email with feedback from an informal observation, they sent a response addressing any questions or concerns. Significantly, however, teachers sometimes punctuated or temporarily interrupted the usual status quo when they accepted particular opportunities to interact their responses by adapting their response to one kind of demand for use or application in their response to the other. As discussed below, punctuated compartmentalization obtained across all the focus PLCs and focus teachers, though with variation in the opportunities for punctuation and the types of adaptation.

"Compartmentalization" characterizes the default or baseline state of the relationship between focus teachers' responses to the demands of PLC and evaluation. The interview data were unambiguous here. Ms. Daniels, a Walker focus teacher, initially characterized the

evaluation process and the work of her Math 10 PLC as "separate," but she then elaborated saying:

> They do seem kind of compartmentalized. In one way, your PLC, we're looking at our students, what can we do to get all [Math 10] students to succeed? And then the other way, an evaluation is very individualized. It's just, what does my teaching look like in my classroom? What's my individual impact on the school? So I do think they are a bit compartmentalized.

In Woodford, Ms. Dalmore, similarly characterized evaluation as "its own silo for the most part." When in a later interview Ms. Dalmore was asked whether evaluation "had any effect on how you participate in departmental or grade-level PLC time, what you do with your work or how you look at your work there," she responded, "Honestly, no. I mean, I would be doing everything that I'm doing regardless of whether evaluation was attached to it or not." This conclusion was echoed by focus teachers in the Walker PLCs. Ms. Williams and Ms. Trace (ELA 10) and Ms. Daniels and Mr. Roses (Math 10) explained that their respective PLCs did the same work in the same way regardless of evaluation. Focus teachers at both case schools implicitly expressed a compartmentalization perspective when they said that they experienced little or no direct effect or influence of one on the other in preparing or making responses (Ms. Laphroaig, Mr. Roses, Mr. Bulliet, Ms. Williams, Ms. Trace), and/or that the evaluation process was mostly unrelated or unconnected to PLC work (Ms. Macallen, Mr. Bulliet, Ms. Williams, Ms. Trace, Ms. Forester).

In addition to generally seeing their PLC and evaluation as largely separate responses to separate demands in timing, content and nature, the same sort of relationship emerged from other patterns in the interview data. Teachers in each focus PLC described how their evaluation responses tended to be narrowly focused in time and content whenever a specific evaluation

demand was imminent. Ms. Dalmore, referring to her colleagues at Woodford more broadly, elaborated on her thinking of evaluation as siloed this way:

> I mean, people tend not to think about it or worry about it until, you know, when the data is due, or they're being observed, or they're having their preconference, or something like that. Like, it feels like a zoom-in snapshot of this one moment in time and then maybe again later in the semester. But it doesn't feel like an on-going thing.

Her ELA 11 colleague Ms. Macallen indicated that she awaits prompts from her evaluator, a very much "check the boxes kind of person," to which she gives the particular response she believes he wants. At Walker, Ms. Daniels averred that, even as a probationary teacher, she placed little emphasis on evaluation, changing little of what she usually did for any of her observations. Mr. Roses, in referring to several components of evaluation, explained that

> … I don't really think about it [evaluation] very much at all. I mean, it's not on my mind until there's something that I have to do, you know, until it's the day we were supposed to pick our goal … Or like the mid-year review, I don't really worry about it. I didn't worry about it until, you know, a day or two before I kind of thought about what I was going to say or what I wanted to talk about. … Same thing for the evaluation [formal observation]. If they come to my classroom to observe me, you know, it's not anything I really worry about until that happens.

Finally, among Walker ELA 10 focus teachers, Ms. Forester stressed that she gave little thought to evaluation, and Ms. Williams felt like "evaluation is just something that happens."

In a similar vein, it seems that evaluation demands or teacher responses were rarely, if ever, raised in any significant way during PLC work or in connection with PLC demands or responses. In a follow-up interview, Ms. Daniels explained further how the Math 10 PLC worked regardless of evaluation:

> … I don't really see too much overlap [between PLC and evaluation]. Like often in our PLC group, we're not even worried about evaluation or what administrators may think,

> or how this will affect our end-of-year effectiveness forward. We're just honestly going about, you know, day-by-day, week-by-week, just teaching our content the best we can. Not really thinking of the evaluation piece.

Mr. Roses and Mr. Bulliet independently agreed that evaluation was not a subject of PLC discussion, mainly because it was not relevant to more pressing PLC work, which directly contributed to their practice. With respect to the ELA 10 PLC, only Ms. Williams, a probationary teacher, said she had asked her PLC colleagues their view about certain evaluator comments during prior evaluations. She did not give a specific example; but as she described them, they sounded like informal interactions in which Ms. Williams sought her colleagues' opinions about what her evaluator may have meant. These conversations in any event did not appear related to the PLC or its work. No one else suggested otherwise, nor that other discussions about evaluation occurred in any relevant sense.

According to Ms. Dalmore at Woodford, by the 2019-20 school year, evaluation was not something that teachers brought up in PLC meetings or that pertained to PLC work. Aspects of evaluation had been an intermittent item of teacher talk during PLC time in the initial years following the district's adoption the Marzano evaluation tool. The decision precipitated significant (and unpopular) changes in evaluation and PLCs (see Chapter 6). Similar to other changes in the evaluation system, some teachers in response went to their PLC colleagues for understanding and support with the process, requirements, and stress. More specifically, Ms. Dalmore said that teachers sought help and assurance with having the required Marzano scale (a kind of learning progression rubric for each curricular standard in a unit) and related material set for their observations. It was in this connection that Ms. Dalmore mentioned an occasion when a scale one teacher created for her evaluation was later used in her PLC to discuss instruction in the relevant unit. This is the sole example of an occasion when a teacher's evaluation response

directly affected a PLC's work. Ms. Laphroaig, too, described several occasions when she informed her PLC colleagues about the completion of evaluation system forms, administrative evaluation expectations, and lessons and materials she had used successfully for observation.

All these sporadic instances seemed to have occurred in passing; they were certainly all past. They happened while the PLC/evaluation changes due to the Marzano framework were being implemented, and they were unrelated to the principal ELA PLC work assigned for the 2019-2020 school year (see Chapter 6). The ELA grade-level/content PLCs had previously identified "priority standards" that they had incorporated into the Marzano scales they created. These, along with common assessments, had been aligned with the instructional units. As a part of my observations during the year in which this study took place, the ELA 11 PLC first reviewed the completion of this prior work, which was to form the basis for a course pack of materials. It then proceeded with the next part of a course pack—namely, drafting a narrative summary of the course and each of its units. In any event, like informal side conversations among teachers generally, these prior PLC/evaluation connections were at most ancillary to the ongoing work, objectives, or functioning of the ELA 11 PLC. More important, although some may have had their source in evaluation, they were—and most certainly are—largely unrelated to any PLC demand or response.

If the compartmentalization of responses represents a sort of general rule, then each punctuation denotes an exception. Punctuation occurs when, given an opportunity, a teacher "de-compartmentalizes" their PLC and evaluation responses, and instead opts to connect them for a limited time and particular purpose. More precisely, punctuations are occasions of agentic, temporary, and goal-oriented departures from the usual compartmentalized relationship between

responses to the opposite state: the interaction of responses such that the response to one demand is adapted to respond to the other.

Even as focus teachers indicated that day-to-day compartmentalization was in effect "business as usual," each focus PLC's teacher-members readily identified the same, apparently highly salient, example of punctuation—i.e., a past or present PLC task or product that could be re-purposed for responding to an evaluation demand. All ELA 11 PLC study participants repeatedly pointed to the Marzano scales on which ELA PLCs had been working since the prior school year, and which administration required for the observation component of evaluation. All teachers in the ELA 10 PLC spoke of how their evaluators consistently solicited information about their PLC work during evaluation conferences. And the Math 10 PLC was united in determining to work on the collective analysis of individual teacher-member student growth data as also required for evaluation. While indicating compartmentalization by describing PLC and evaluation as "removed from each other" and "in two separate realms," Mr. Bulliet also aptly characterized punctuation when in the student data analysis context he explained:

> I definitely think that they do have a hand-in-hand relationship in some respects, like because we're supposed to gather data for both of them. … Some of it overlaps, and some of them like do kind of work together.

On the district and school levels, opportunities for punctuation had their source in administrator structured and implemented relationships between evaluation and PLC. (The next chapter addresses how such structural relationships, as well as their interactions, shaped teachers' adaptive responses.) From a teacher's perspective, such opportunities actually arose in the timing and substance of the individualized evaluation demands to which they were responding. Teachers could apply certain PLC work in response to certain evaluation demands (for how, see the next section), and thereby benefit their own interests (for why, see the last section of this

chapter). Conversely, however, focus teachers saw little or nothing about evaluation that seemed timely or pertinent, let alone a benefit, to what their PLCs did or how their PLCs worked. Given the structural prompts to punctuation, together with the asymmetry in the useful interaction of responses, focus teachers pointed adaptations of responses in one direction only. All focus teachers recruited certain PLC responses to support of their evaluation responses, and not the other way around.

Consistent with focus teachers' accounts of their PLC work, the observational and documentary data reinforce the finding that PLC work was actually conducted in response to PLC demands and irrespective of evaluation (even if sometimes PLC responses were later adapted for evaluation). Further, when compartmentalization was punctuated, past or ongoing PLC responses could (and did) contribute to teachers' evaluation responses, but not vice versa. Without exception, observations showed that all focus PLC work concerned PLC demands. Indeed, with the sole exception of a passing comment during a Woodford ELA 11 PLC meeting, *none* of the observational data includes *any* mention of teacher evaluation, even in general terms, let alone an aspect of any teacher's response to evaluation. This absence of mention included the periods of time when teachers informally interacted before and after the regular meeting time of the focus grade-level/content PLCs. In contrast, as shown in more detail below, participating administrators and focus teachers agreed that PLC work and products often arose at particular points in the evaluation. Moreover, again as the following chapter section explains, teachers in their evaluation responses could (and did) use or refer to PLC products—documents or writings created for PLC purposes—without reference to evaluation. There is no such evidence that the reverse occurred.

All focus teachers and PLCs at both case schools acted within an overall context of punctuated compartmentalization. Most of the time teacher responses to PLC and evaluation demands were independent, self-contained. Yet, within the structural relationship between evaluation and PLC, as differently implemented in each case school, a range of opportunities for punctuation prompted focus teachers to interact responses in a particular way: collective PLC responses to PLC demands were adapted to help make individual responses to evaluation demands. How this process occurred, including the emergent patterns among adaptive response variations, is the subject of the next section.

How Punctuation Happens: Interaction Opportunities and Adaptive Response Strategies

Focus teachers punctuated compartmentalization within the structural relationship of evaluation and PLC, when an opportunity arose that they could act upon it in a timely and beneficial manner by adapting an available PLC response—i.e., past, present or potential PLC work product—to a present or impending evaluation demand. To work successfully, the process depended on its operative elements—structural opportunity and agentic action—fitting together like a key in a lock. Stating and supporting my findings about the punctuation process requires the identification and analysis of the kinds of locks and corresponding keys. Although both elements varied within and across cases, how focus teachers adapted certain PLC responses to satisfy certain evaluation demands forms a useful typology of teacher adaptive response strategies, at least in the context of the case schools.

Some of the opportunities were explicitly communicated through related PLC/evaluation structures. Thus, for more than a year before 2019-2020 the Woodford district had tasked grade-level/content PLCs like ELA 11 with creating Marzano scales, together with aligned common assessments, for the standards applicable to each curricular unit. The applicable standards

themselves were those from the Common Core State Standards (as adopted in Michigan, the MI CCSS). At administration direction, the Woodford ELA 11 PLC had previously designated some of the MI CCSS as "priority standards" for the ELA 11 curriculum, assigning them to the ELA 11 units in which they were covered and assessed. For formal observations, teachers were required to submit lesson plans that included the relevant scale(s), with embedded priority standard(s). The lesson documents and materials were then reviewed in the pre-observation conference, and they could also be a subject of discussion at the post-observation meeting. The common, standards-aligned assessments were used apart from the observation component to collect data for a portion of the separate student growth component of evaluation.

It was hardly surprising, then, that the focus teacher-members of the ELA 11 PLC stressed how they and their colleagues made sure to import a scale from the prior focus PLC work into the observation component of their evaluation in subsequent years. Despite differing views on the instructional utility of the scales, Ms. Dalmore, Ms. Laphroaig, and Ms. Macallen all responded without hesitation to this evaluation demand with a specific PLC product that lay right at hand. Even Ms. Macallen, the most negatively disposed toward both evaluation and PLC demands, admitted her appreciation that her PLC had created scales. "I could easily plug [a scale] into my evaluation tool. Had I, had we, not already done that work, I would have been fairly clueless about how to create those on my own." Similarly, when each Woodford focus teacher selected the common assessments with which to collect student growth data, she was again re-purposing for evaluation a PLC product created in response to a prior PLC demand.

In Walker, as in Woodford, explicit structural relationships between PLC and evaluation demands provided opportunities for adaptive responses, although between and within case differences emerged in how structures interacted and teachers responded. For example, Walker

teachers were not required to submit lesson plans or related materials in preparation for their formal observations. Still, lesson planning activities were a part of the first domain of the Danielson evaluation tool in Walker. Teachers could voluntarily submit plans and materials as supporting artifacts, and administrators could review and discuss them in the pre- or post-observation conference. The Walker focus PLCs engaged in some form of common lesson planning. The ELA 10 PLC primarily concerned itself with lesson sequence, resources, and formative assessments. The Math 10 PLC took a more comprehensive approach by preparing and revising complete common plans specifying every element of the lesson. In both cases, focus teachers adapted aspects of their PLC's work for use during the observation component of their individual evaluations.

Under the Walker evaluation system, fifteen percent of a teacher's effectiveness rating consisted of setting and achieving at least one approved "professional performance goal" over the course of each year. Teachers at Walker High School could respond to this individual evaluation demand by voluntary agreeing with their colleagues to a departmental or grade-level/content PLC goal. Once approved by building administration, the collective goal was a PLC demand to which PLC work during the year responded. At the same time, each of the PLC's teacher-members was responding to the goal component of their individual evaluations. For the 2019-2020 school year, both the ELA and Math Departments, and thereby their grade-level/content PLCs like Math 10 and ELA 10, made goals concerning curricular standards and aligned assessments. Despite these goals' common subject matter, the PLCs with their respective focus teachers, ended up adapting responses in different ways.

The Math 10 PLC had already written its "essential standards" and common assessments. These were aligned with the MI CCSS and SAT, but they were fewer and more focused on what

the PLC's focus teachers had determined constituted the most important student learning. The PLC therefore began to the revise them, along with some of the common lesson plans they had prepared and piloted in 2019-2020 when they had revamped the entire curriculum. After the 2019-2020 year began, the Walker central administration decided that the Math Department (among others) should create a curriculum handbook or guide, mainly for students and parents, that would for each math class set forth its essential standards or learning targets, and give sample assessments. The Math 10 PLC folded this task into the work in which it was already engaged. Completing the district document became the outcome satisfying the departmental PLC goal, and *ipso facto* one of the teachers' individual evaluation goals.

The ELA 10 PLC, in contrast, was not at the point of a wholesale review of its curriculum. (It was to be the PLC goal for 2020-2021.) This PLC relied on selected MI CCSS tied to previously established units, although in the last few years, it began creating common assessments. The preceding school year a district curriculum committee that included Ms. Trace and the Walker ELA department chair had begun to develop a curriculum guide or packet to vertically align ELA standards district-wide. By the 2019-2020 school year, the district expectation for this internal document was expanded such that the curricular standards for each high school course would be aligned with newly created assessments and rubrics. These would be created in a standardized form for use solely to determine student growth in a class during each trimester. The ELA department adopted the completion of the guide as its PLC goal, which meant it could become an evaluation goal of the individual ELA teachers as well. Since Ms. Trace continued to lead the effort, this goal became the one that counted toward her evaluation.

Both focus PLCs had in prior years created common assessments, on which they continued to work during the 2019-2020 school year. An essential demand of the DuFour PLC

model, which Walker building administrators emphasized, is that teachers in a grade-level/content PLCs must collect, analyze and compare their students' common assessment performance data as the basis for collaboration around curriculum, instruction and assessment. However, it appeared that few Walker PLCs had yet fully or consistently adopted the DuFour process. Moreover, to the extent student performance data were reviewed, the Walker focus teachers, whether in their PLCs or otherwise, could individually select the classes and common assessments to analyze, share and compare.

The district relied on school-level NWEA and SAT test data to show student growth for that component of individual teachers' evaluations. In September 2019, as one of several changes in the evaluation process, Walker central administration for the first time specified levels of student performance in each teacher's classes that directly contributed to the teacher's student growth effectiveness rating. To respond to this new evaluation demand, teachers had to change how they collected, analyzed and reported their student growth data.

Although the ELA and Math departments adopted a common understanding of how to compute a student growth effectiveness level, the Walker PLCs differed in their responses. The focus teachers in the ELA 10 PLC planned to change their individual evaluation responses to accommodate the new demand, but their collective PLC work would remain unaltered. The Math 10 PLC, in contrast, saw an opportunity to at once meet the new evaluation demand and the existing PLC demand to collect, analyze and discuss individual student data using the same lesson based on the same standard with the same assessment. The focus teachers accordingly agreed to a few essential standards with aligned common assessments that would be taught and given at virtually the same time during the second trimester. Then, as a PLC, they planned to analyze the collected data as a PLC, both for PLC and evaluation purposes. With time and

practice, this process of "killing two birds with one stone" could be effectively reiterated each trimester going forward. As Ms. Daniels summarized the PLC's data analysis process, "it'll be helpful for us to compare data, plus for us to present data for end of year evaluation."

Not all punctuation opportunities and adaptive responses were as directly and specifically related as those that involved the use of common PLC products (Marzano scales, standards, assessments), or the concurrent work in response to an overlapping PLC/evaluation goal or data analysis demands. Some instead depended on more open or flexible opportunities for teacher initiative or exploitation, where demands were less narrow or constraining, and evaluation responses more indirectly invoked PLC work more generally. Most prominently in Walker and Woodford, punctuation opportunities arose during evaluation-related meetings and similar back-and-forth communications between teachers and administrators. These personal interactions presented relatively frequent opportunities to favorably adapt or characterize PLC work and products to support evaluation responses and results. Focus teachers nonetheless varied in their interpretation of any given opportunity and so in their agentic response.

The Walker district first mandated individualized "mid-year review" meetings for the 2019-2020 school year. Neither the specific purpose of the meetings, nor the impact on evaluation were clear to teachers or building administrators who were to arrange and conduct them. Absent guidance, the Walker principal, Mr. Creek, decided to structure the meetings around his favored school initiatives. He sent each teacher a brief nine question "survey" that asked about such matters as teachers' progress toward their yearly goals; their PLC's work, especially concerning student data; how they were planning comply with the new student growth standard; and any suggestions for improvement in school initiatives, including PLCs. The

implication was that the actual mid-year review meeting with Mr. Creek would take up the same questions.

Focus teachers could respond to this new evaluation demand at least in part by adapting PLC responses, particularly since Mr. Creek had deliberately solicited teacher advocacy of their PLC work. Yet, focus teachers differed in taking the opportunity. On the one hand, Ms. Daniels' response extolled the successes of the Math 10 PLC, as well as the PLC's plan to analyze student data. (The closure of schools due to the Covid-19 pandemic prevented me from obtaining either Mr. Bulliet's or Mr. Rose's questionnaire. Nor did I have the opportunity to interview Mr. Bulliet about his mid-year review meeting. Both were in other contexts strong advocates for PLCs generally and their PLC specifically.) On the other hand, the focus teachers in the ELA 10 PLC mainly gave cursory responses that did not promote their PLC, its work, or their own contributions.

The ELA focus teachers described their individual meetings with Mr. Creek in similarly prosaic terms. According to Ms. Williams, for instance, the survey "just kind of felt like general questions." At the meeting, she and Mr. Creek "talked about the [curriculum document] that we're doing for central office. We talked about how the PLC is going. Talked a little bit about stuff for next year, as far as reworking classes. That sort of thing. … So it just kind of felt like a nice little check in to make sure we're all on the same page and right direction." Ms. Trace spoke in general terms as well: "We just talked about things that we can do to improve the school, and you know, things that we think are going well and things that collectively need improvement." Mr. Creek "asked questions about how well we're getting along, how well, you know, we've progressed in this latest goal." For these focus teachers, the mid-year review seemed a required

but more or less insignificant step they needed to complete. Ms. Forester summarized the process thus:

> Well, we did, for the first time this year, a mid-year evaluation questionnaire before we went into speak with our administrator. So that was e-mailed to us. And then there were some talking points for getting into them online. And then we went in to schedule a time to meet. And then we just kind of went over how things are going, schoolwide, classroom wide, PLC—all of those types of things.

The way the Math 10 PLC focus teachers approached the same punctuation opportunity differed from that of the ELA 10 teachers, even as Ms. Daniels' experience diverged from that of her PLC colleague Mr. Roses. Because the Walker assistant principal, Mr. Mark, unexpectantly went on a leave, all math teachers met with the district superintendent instead of Mr. Creek. Ms. Daniels, the first to have her mid-year review meeting, came away incensed because the superintendent had little knowledge about her teaching and other professional work, for example as a coach. Worse, "we did not discuss my PLC work in any way." She said the meeting was pleasant enough, but "I honestly feel like I got nothing from the meeting. …I basically just heard the superintendent's vision for the next year and the things he would like to implement. Totally different from what I was expecting." Mr. Roses, expressed a diametrically opposed response to his meeting:

> …I kind of looked at it as an opportunity. I don't know that I've ever had a conversation with [the superintendent] other than when I was hired. … So it was good to sit down with him, and kind of let them know what I'd been doing this year and what our PLCs had been doing, and things I'm excited about and to kind of get some positive feedback from him.

Mr. Roses, in other words, took the initiative in his response to an evaluation demand to promote what the Math 10 PLC had accomplished, his contribution to it, and its positive effect on his teaching.

At both case schools, observation-related and year-end evaluation conferences sometimes provided opportunities for focus teachers to adapt for evaluation their PLC's work, products and results, as well as their individual participation, improvement, and contribution to district/school goals. These response interactions and adaptations took two forms: first, focus teachers could advocate for their PLC and their role in it; second, focus teachers could rely on their PLC to explain, justify, or project confidence or faith in their choices during the evaluation process. Again, however, there was between and within case variation in how any adaptation occurred.

Despite her reaction to her mid-year review meeting, Ms. Daniels had used evaluation meetings both to advocate for and to rely on the work of the Math 10 PLC. According to all its members, their PLC jelled during a highly productive 2018-2019 school year. Ms. Daniels explained that, when the PLC's focus teachers went into their year-end evaluation meetings, "we were excited to share the work we had done," and so each did. The PLC's effective collaboration continued into 2019-2020, which in part set her up for the disappointing mid-year review meeting. Going in, Ms. Daniels "wish[ed] that my PLC work was helping to, like, guide the evaluation piece" as it had in 2018-2019. By the end, she had concluded, "but that was not the case for my mid-year review." Still, in connection with the observation component of evaluation, Ms. Daniels emphasized she could rely on her PLC work to support her response to the evaluation observation demand:

> I would say the fact that I get to do, like, cool and engaging activities when my administrators observe me that stemmed from PLC. I didn't come up with a lot of those myself. We came up with those together.
>
> So we definitely put our brains together, came up with creative things, not necessarily for observation's sake, but we wanted to have a few really cool things that we did in our classrooms each trimester… .

Mr. Bulliet referred in general terms to the same sort of collective confidence and validation the Math 10 PLC provided him, which strongly supported his professional growth and thereby his response to evaluation demands. Mr. Roses even more specifically tied the collective quality of the Math 10 PLC's to his evaluation performance:

> I feel like, by working together, that what we do on a daily basis is better. And because I know that it's, that the instruction is good, I know that the assessments are good, that we have a wide variety of experiences for the students, but I'm not really concerned about when I get observed.

All the focus teachers in the ELA PLC acknowledged opportunities during their evaluation meetings to promote or rely on their PLC work to support their response to individual evaluation demands. Like their responses to the mid-year review meetings, however, they were vague about the discussion. Ms. Trace said that the topics of her evaluation meetings and her PLC work during the first half of 2019-2020 "were pretty intertwined because a lot of things that I brought up in my evaluation were things we were working on in [the ELA 10 PLC]." Yet, the examples Ms. Trace mentioned were limited to the district curriculum document, which did concern the interaction of Ms. Trace's PLC and individual evaluation goals, and a proposal for a new college preparation (honors) level ELA 10 course, which was unrelated to evaluation. Ms. Williams merely mentioned that during her post-observation conference, "it's brought up, you know, how is your PLC time? How are you working on that?" Referring to her evaluation meetings, Ms. Forester noted that, "[w]ell, we always discuss PLC because our administrator [Mr. Creek] really likes PLCs [and not evaluation.] … So, I mean, yes, we'll definitely talk about PLC and PLC time." Ms. Forester for her part suggested that PLCs would be more effective, for example in analyzing student data, if departments were not assigned so many administrative tasks during PLC time.

It seems that teachers in Woodford had opportunities to present their PLC work as supporting their observed lesson because administrators introduced or invited the subject during evaluation meetings (see Chapter 6 for the goal and design orientations underlying the administration's approach). According to Ms. Lauder, the secondary education director, building administrators such as Mr. Regal and Dr. Sark were having "very intentional conversations" during pre-observation conferences about the teacher's lesson and the Marzano tool. These evaluation discussions related back to the teacher's PLC work. They indicated "where teachers are in their thinking on the standards-based model and using a scale with kids, and being comfortable with what it means to move through the scale, which is the work of the departments." The Woodford High School principal, Mr. Regal, answered "absolutely" when asked whether a teacher's participation in their department time PLCs ever arose as an issue in evaluation meetings. "We talk a lot about how they've worked with their other colleagues to develop this unit or gotten feedback about things or how the department [in grade-level/content PLCs] decided on this scale or this progression. So, yes, it [PLC work] definitely comes up during those pre- and post-[observation] conferences." Dr. Sark, an assistant principal, spoke of pre- or post-observation meetings when a teacher would explain a lesson in part by referring to their PLC discussions about it. (Dr. Sark did go on to point out that, while she appreciated teachers' attempts to engage in such PLCs conversations, she was disappointed that they didn't concern the Marzano tool or its instructional strategies.)

If the administration's perspective seems clear enough, the extent of teachers' adaptive responses to the opportunities presented is much less so. Ms. Macallen, who found "insulting" both administration-driven PLC tasks and evaluation processes, did not identify any evaluation conversation involving her PLC work or even her use of PLC products in evaluation. With some

prior familiarity with the Marzano framework, Ms. Dalmore described how she modified her lessons for her own observations and how she incorporated PLC work products. Like Ms. Macallen, however, Ms. Dalmore did not relate any evaluation discussion about any PLC/evaluation connection. Ms. Dalmore did indicate that other teachers' PLC work may have indirectly benefited them in evaluation:

> I guess there's a sense of confidence that comes from PLC that helps our teachers in terms of their actual evaluation. You know, feelings like, hey, I'm doing things that I need to do. Here's what my department and I have talked about, I know I'm teaching the right thing.

Ms. Laphroaig in her evaluation conversations made explicit an affirmative PLC/evaluation interaction. She described her observation preparation and related lesson discussion this way:

> I essentially pulled that scale for [the relevant unit]. And I looked at it, looking at the prompts, looking at the thesis, what the scale requires. I pulled that piece out and added it to my lesson plan, so I could show my administrator, this is the scale that I've taken this from; this is the piece I'm working on today; this what you're going to see. So I was able to take that work that I've been doing within my [PLC] for the last couple of years and then apply it directly to my evaluation by saying, OK, this is the scale that we developed, and this is what I'm focusing on, and this is how I'm going to get there.

The interview, observational, and documentary data all converge on the identification, summarized above, of those instances when and how focus teachers in focus PLCs punctuated the default state of evaluation/PLC compartmentalization. The analysis of what teachers and PLCs did to interact their responses reveals four meaningful categories along two intersecting dimensions that represent agentic patterns or forms of how punctuation occurred—i.e., the ways or strategies teachers employed to adapt PLC work done in response to PLC demands for application in evaluation responses to evaluation demands. More specifically, I propose four types of adaptive response strategies, four distinguishable kinds of keys—teacher agentic

responses—that fit the locks—the punctuation opportunities presented by the structural evaluation/PLC relationships in the case schools. I have termed the four proposed types of strategy (1) Selective Use, (2) Dual Process, (3) Professional Validation, and (4) Lobbying. Each is descriptively defined with generic examples in Table 5.1.

Table 5.1 Teacher Adaptive Response Strategies

Adaptive Response Strategy	***Description***	***Examples***
Selective Use	Teachers' collaborative work in response to a PLC demand results in a product that may later be used, adapted for use, or be a model for use in response to an individual performance evaluation demand.	• To plan for and conduct a lesson for formal observation, a teacher incorporates into their lesson plan and materials the curricular standard, rubric, and assessment that their PLC produced. • To collect and analyze student growth data, a teacher uses the common assessment that their PLC produced.
Dual Process	Teachers' collaborative work in response to a PLC demand involves a goal, task or activity that, together with the end product, simultaneously constitutes a response to an individual performance evaluation demand.	• A PLC's teachers use their PLC-produced curriculum standard, assessment, and scoring method for collecting and analyzing student data both to show student growth for evaluation purposes and as a basis for PLC collaboration over instructional strategies and formative assessment. • A PLC-selected collective goal is also selected by a PLC member as an individual evaluation goal; work on and the achievement of the same goal satisfies both a PLC and an evaluation demand.
Professional Validation	Teachers' collaborative work in response to a PLC demand supports a response to an individual performance evaluation demand with peer-created, approved or endorsed product.	• A teacher relying on PLC-produced and piloted lessons feels justified or confident in instructional choices and so better prepared for observation, and can better respond to evaluator ratings, questions, and feedback. • During a pre- or post-observation conference, a teacher justifies their lesson choice and content based on PLC-produced lesson components and materials.
Lobbying	Teachers' collaborative work in response to a PLC demand provides evidence of individual contribution and/or PLC success to support or promote a response to an individual performance evaluation demand.	• A teacher uses a mid- or year-end evaluation conference to advocate for greater consideration of PLC work (obstacles faced and/or accomplishments) as affecting individual performance and/or to show furtherance of administrative goals. • In response to evaluation-mandated reflection, questions or evaluation tool domain, a teacher advocates for value of PLC work and contributions; includes the submission of PLC artifacts.

The strategies are not mutually exclusive. Teacher agentic responses may involve more than one based on how a given teacher sees and takes opportunities for punctuation. For example, for her observation lesson preparation Ms. Laphroaig selectively used a Marzano scale, which she then professionally validated during her pre-observation conference. Ms. Laphroaig's ELA 11 PLC colleagues, Ms. Macallen and Ms. Dalmore, had the same opportunities in the evaluation process, but it seems they applied the Selective Use strategy alone. Similarly, the Math 10 PLC focus teachers together selected a few essential standards and aligned common assessments to use for student data collection and analysis. They planned, in addition, to take advantage of the Dual Process strategy, as they would in their PLC examine the student data for evaluation and PLC purposes. At least before the pandemic, each ELA 10 focus teacher intended to selectively use common assessments for individual computations of student growth.

The four adaptive strategies can be meaningfully compared and distinguished along two intersecting dimensions. One of the dimensions contrasts specific direct uses with holistic indirect uses of PLC responses (work and/or products). Selective Use and Dual Process represent the former because teachers can directly satisfy to an evaluation demand with a specific PLC product. They are distinguished along this dimension by the timing of production: Selective use involves production in response to a past PLC demand and before use in evaluation, whereas Dual process involves production and use in evaluation through the same process as responds to a contemporary PLC demand. Selective Use occurred when Woodford focus teachers plugged a portion of an existing Marzano scale into the lesson plan for their formal observation and when Walker teachers selected the PLC-created common assessments with which to collect student data. Dual Process occurred when Walker Math 10 focus teachers collected and analyzed student data for both PLC and evaluation purposes, and when Ms. Trace worked on the curriculum document, which simultaneously moved her toward the achievement

of her evaluation and PLC goals. Holistic Professional Validation and Lobbying involve holistic indirect uses because teachers' response to an evaluation demand advocates for or relies on the collective authority or professional bona fides of the PLC *qua* PLC. As a pair of holistic indirect strategies, Lobbying and Professional Validation differ in their "offensive: as opposed to "defensive" orientation. The former occurred, for example, when Ms. Daniels and Mr. Roses in evaluation meetings affirmatively promoted the successes of the PLC they effectively contributed to. The latter was shown when Ms. Laphroaig justified her lesson plan in an observation pre-conference with reference to its implicit PLC endorsement. Holistic indirect uses rely on a teacher's proactive choices to employ them, while teachers' decisions for specific direct uses are more reactive, requiring a close match in PLC/evaluation demands.

The intersecting dimension compares the relative degree of teacher versus PLC focus of the adaptation. Selective Use and Professional Validation both entail an adaptation focused on what the PLC collectively did or made available—created the product for use or the record for support in evaluation responses. Put another way, the adaptation could not occur unless the PLC had first appropriately responded to PLC demands. The high teacher/low PLC focus strategies—Dual Process and Lobbying—also involve PLC responses to PLC demands; but in each case, the focus of the adaptation lies in what or how the teacher contributed to the PLC's work. Even when the Math 10 PLC implemented the Dual Process adaptive response strategy, each teacher still was individually responsible for contributing their data, participating in its analysis, and reporting effective results for the personal evaluations. Similarly, the focus was on Ms. Trace's role and goal when she in an evaluation meeting explained the challenges she faced as the leader of her department PLC in completing the district curriculum document.

Figure 5.1 places each of the four adaptive response strategies in the appropriate cell representing the intersection of the two dimensions.

Figure 5.1 Two Intersecting Dimensions of Proposed Typology

Degree of Teacher/PLC Focus of Adaptation	***Specific Direct Use*** **(More Demand Dependent)**	***Holistic Indirect Use*** **(More Response Dependent)**
Lower/Higher	Selective Use	Professional Validation
Higher/Lower	Dual Process	Lobbying

This section has shown that, during the individual evaluation process, each focus teacher took at least one opportunity provided by related PLC/evaluation demands to adapt their PLC's work, which constituted collective responses to PLC demands, for use in responding to evaluation demands. To account for the totality of teacher adaptive responses, this section has also proposed a typology of four adaptive response strategies characterized by two cross-cutting dimensions. The final section of this chapter completes this study's findings as to the first research question by examining the principal factors motivating teachers' agentic deployment of adaptive response strategies in the structural contexts of the case schools.

Why Teachers Punctuate Compartmentalization: Teacher Goals and Response Values

This study finds both that focus teachers typically responded to evaluation and PLC demands separately as they independently arose and that teachers sometimes punctuated the usual state with adaptive response strategies. Why did teachers engage in affirmative departures from the status quo? This section shows that focus teachers were motivated to punctuate upon the alignment of their goals in responding to PLC and evaluation demands with the relative value

or benefit they derived from engaging with each. Further, the same correspondence of goals and values helps explain the compartmentalization of responses in the first place.

Focus teachers derived little benefit from the routinized evaluation systems in Woodford and Walker. The low value of evaluation was directly and reciprocally tied to teachers' principal goal in responding to evaluation demands—compliance. Practically every focus teacher in one breadth both disparaged the evaluation process and expressed resignation to its mandate.

At Walker, Ms. Williams characterized evaluation as "a lot of busy work," but also as "a necessary evil that somebody else is making us do." Ms. Williams' ELA 10 PLC colleague, Ms. Forester, succinctly agreed: "[F]or me at least, it's just, it's something that has to be done, but it doesn't really mean much to me." One of the Math 10 PLC focus teachers, Mr. Roses, similarly explained that "evaluations are not something that has ever done anything good for me as a teacher. It's just something that I've had to do." Mr. Bulliet supplied a final example from the Walker focus teachers, when referring to the observation component, he said: "So I don't feel a ton of usefulness from it. It's kind of a hoop that I have to jump through because it's part of my evaluation[.]"

The Woodford focus teachers all expressed similar sentiments in almost identical terms. Every one used the "hoop" metaphor. Ms. Dalmore used it twice, including when she said, "It's like, oh, it's just one more hoop to jump through, because it doesn't feel like, oh, these are things I want to do." Ms. Macallen mixed three metaphors in her characterization of the observation component, calling it "all busy work and just having to jump through those hoops and check the boxes, and get it over with so they can add it to our overall evaluation score at the end of the year." At another point, Ms. Macallen summed up the evaluation process as a whole:

> [I]t's just another thing we have to do throughout the year. …
> [M]ost people I talk to, whether they're new teachers or have

> been here as long as I have or longer, they find it equally tedious and just a hoop they have to jump through and get it over with.

Focus teachers offered various, often overlapping rationales for their low estimation of the value of evaluation and of the purpose of responding to its demands. Often differences among responses turned on the emphasis given one or another reason. Math 10 PLC colleagues Ms. Daniels and Mr. Roses agreed that evaluation was, in the latter's words, "irrelevant to what we do on a daily basis." But at the same time, while Ms. Daniels did not find the process overly stressful, she felt frustrated by the absence of substantive feedback, even when she sought it. Her experience with mid-year review (discussed in the chapter section above) caused her to find the evaluation process "even more useless, because I felt that it was honestly a waste of my time[.]" Mr. Roses and the third Math 10 PLC teacher-member, Mr. Bulliet, did point to the inherent discomfort, even fear, the evaluation process provoked. The focus teacher members of the ELA 10 PLC evinced a similar pattern of interlocking explanations. At one point, Ms. Trace was irked because the observation domain ratings seemed more or less random or arbitrary. At another, she, later joined by Ms. Williams, expressed confidence that, wholly apart from the evaluation process, her building administrators would let a teacher know if their job was in jeopardy; and by the same token, should administrators wanted to fire someone, they would find a way to do so. Ms. Forester's opinion was somewhat akin to her PLC colleagues, as she felt in effect that "no news is good news." She and Ms. Trace in addition maintained that evaluation did not concern them because their building administrators had communicated that they did not find the process particularly necessary or helpful.

Within the overall consensus and supporting rationales, focus teachers' assessments varied somewhat due to idiosyncratic factors. Several focus teachers tried to put something of a positive spin on the process. Ms. Dalmore thought her evaluation was important; it did have

some value, she commented at one point. It provided the impetus to improve otherwise old or stale lessons. She could then present them in a fresh way and receive some feedback. Ms. Laphroaig, a probationary teacher, thought evaluations using the Marzano tool provided a clear and standardized language with which to have conversations. Moreover, "given that it's not something that's optional, I have worked to try to make it, if not an overly valuable experience, at least something that's not a complete waste of my time." For Ms. Laphroaig, the process had resulted in highly effective ratings. Ms. Williams, who had a highly negative experience teaching in another state, found Walker evaluation meetings informal and low pressure, actually affirming in a way. They represented a "kind of confirmation of what's happening," and "it's nice to know we're pretty much on the same page, I think."

Other focus teachers had personal reasons for panning the evaluation process. In Woodford, Ms. Macallen detested the Marzano framework, not only for its adoption as evaluation tool, but also for how its adoption precipitated a change in PLC tasks and expectations. Although she said she maintained a professional relationship with her evaluator, she found it insulting that a veteran teacher like herself was evaluated by an assistant principal who had no K-12 teaching experience. For years at Walker, it was a common belief among teachers that virtually everyone who complied with the evaluation process would be rated "effective." The supposed lack of teachers rated "ineffective" or "highly effective" evidenced the even-handed implementation of evaluation without unfair favoritism or bias. During the 2018-2019 school year, ELA teachers discovered that a significant number of teachers had in fact achieved highly effective status. The revelation especially upset Ms. Trace and Ms. Forester, among other teachers (they asserted). They thereafter expressed distrust of building administration and distain for evaluation, with its results being characterized as "meaningless"

and “worthless.” Ms. Trace’s suspicions grew further in 2019-2020 when the central administration issued new evaluation requirements around the student growth component.

Administrators understood and sometimes shared their teachers’ attitudes toward the low value of evaluation and compliance as its principal goal. Ms. Lauder, the Woodford central administrator, and Mr. Creek, the Walker principal, expressly gave compliance as the principal reason for the implementation of the state-mandated evaluation system. Mr. Creek and his counterpart in Woodford, Mr. Regal, both acknowledged that the severe practical constraints—in time, sampling of performance, subject matter expertise, the press of other duties—on any administration’s capacity to move beyond mere compliance to help teachers improve through meaningful feedback. Most teachers did not seek or expect meaningful feedback; rather, they complied with evaluation demands in exchange for the summative rating. (Some such as Ms. Forester and Ms. Macallen claimed not even to review that.) Dr. Sark lamented the times when she had prepared significant formative feedback, only to have teachers look at the evaluation coversheet with their summative ratings and ignore the rest of the document. “But at the same time,” she admitted, “it is another thing for them to do, just like it’s another thing for us to do.” Given the low value teachers associated with the achievement of compliance, the focus teachers at the case schools appreciated when administration made evaluation “paperwork” less burdensome. The move did not make evaluation more valuable, but it did reduce the cost of compliance.

Not surprisingly, all focus teachers valued their PLCs, especially by comparison to evaluation. Ms. Daniels voiced the common position that “I benefit so much from my work with PLC, and I don’t gain very much for my evaluation.” Ms. Trace made a similar point in somewhat more dramatic language:

> So for PLCs, I would say they're invaluable. I can't imagine functioning without a PLC. I don't know how people do that. And then for the evaluation, I would say it's wasteful.

All focus PLCs exhibited professional interactions, and all focus teachers enjoyed collegial relationships among their focus PLC. Despite such generally shared feelings and the overall positive value focus teachers found in their PLC work, the nature of PLC benefits, and thus the corresponding goals in responding to PLC demands, varied among the focus PLCs. The variation seemed to depend on the degree of PLC autonomy in the selection and purpose of each PLC's work. (Chapter 6 takes up the question of the contrasting design and goal orientations of each case school and district.)

Each school year the Woodford central administration directly assigned specific grade-level/content PLC work and required certain PLC products for completion during the scheduled department or content time meetings. For 2019-2020, the ELA PLCs were to continue working on district project of implementing the Marzano instructional framework through their curricular documents, starting with the scales the PLCs had produced the previous year. Of course, the focus teachers in the ELA 11 PLC differently valued this administration driven work, in large measure depending on their attitude toward the district's purpose (see Chapter 6). But in any event, compliance was the overt PLC goal in the circumstance. Compliance was not the *only* goal, however. Although the PLC did not control what was assigned, it did largely determine how the work was conducted. This allowed for the process itself to benefit teachers. Ms. Laphroaig, who had assumed the role of ELA 11 PLC scribe, explained (and showed) how the PLC could prepare scales and other required curricular documents while still developing a collective understanding of how to maintain desirable control and autonomy in their classrooms. Curricular documents were accordingly drafted without "painting ourselves into a corner where

we felt like we no longer have the freedom to teach in a way that felt authentic to each of us individually." In other words, the PLC could (and did) pursue a goal of compliance with teacher classroom flexibility and autonomy.

Individual focus teachers, moreover, found valuable side effects of the required PLC work, which served other goals such as gaining professional learning and upholding collegial norms. Ms. Dalmore, for example, disliked the tedium of drafting Marzano scales, but she, like Ms. Laphroaig, felt it worthwhile to have something viable to show for their work. In the process, moreover, Ms. Dalmore enjoyed the conversations around curriculum from which she always took something away to improve her practice. As Ms. Macallen put a like view:

> [W]e do end up gleaning some information about different units and books and texts and student writing that is a side piece to—that just comes up naturally in these conversations while we're creating the scales, which I find valuable."

Ms. Macallen stressed that, notwithstanding resistance to working on the Marzano framework, she tried to be "a team player" who understood the need for compliance and so would not leave all the work for her colleagues to do.

The DuFour PLC model in Walker allowed PLCs substantially more autonomy—teacher control over PLC time and work—within broader guidelines and subject to certain expectations for PLC products. The Math 10 PLC had developed into a dedicated and cohesive group that was largely self-motivated and directed. Its focus teachers collectively took responsibility for all aspects of the Math 10 curriculum, instruction and assessment, all with the aim of innovating and improving their own practice and students' learning. Each focus teacher highly valued their PLC work and what they saw as its beneficial effects on themselves and their students. They all spoke, for example, about how their PLC had made teaching more meaningful, exciting, and passionate. The stated and demonstrated goal was continuous professional growth through authentic

collaboration with professional accountability. Mr. Bulliet spoke for the PLC when he reflected that "I feel like a lot of what I do as a teacher has been driven by what we do in our PLC, the way I interact with my colleagues," and when he concluded that "I've definitely grown as a teacher because of my PLC."

The Walker ELA 10 PLC also valued their open and honest collaboration, which mainly concerned preparing for and problem-solving the more practical issues of day-to-day practice. As Ms. Williams summed up the matter:

> I think that one of the most valuable times we can get as professionals is when we get to talk to each other about the specific day-to-day stuff that we do and why we do it. … And so having a place that essentially is a forum to talk about that and be thoughtful is extremely valuable.

For Ms. Trace and Ms. Forester, the Walker's PLC model served to formalize the relationship of mutual support they had formed as new teachers a couple of decades before. The formalized ELA 10 PLC was then available to admit Ms. Williams to the relationship when she later came to Walker. Again, unlike evaluation, ELA 10 focus teachers perceived benefit from PLC work toward PLC goals.

This is not to say that the Walker focus teachers found all PLC work of value, or that no PLC work involved low value compliance. For example, departmental PLCs were required to meet in one place at the beginning of each period of PLC time. There they often had to first address administrative matters such as technology or other materials issues, class scheduling and composition, or other administrative concerns. Teachers were also required to view online presentation and complete review quizzes about the DuFour model of PLCs. Focus teachers to a person did not find these activities particularly beneficial, even if required.

More telling were the directives that Math Department PLCs produce a curriculum handbook and that the ELA Department complete a far more detailed curriculum guide. These

projects entailed a PLC compliance goal, but the two focus PLCs reacted differently in terms of value. Ms. Trace, the most involved ELA 10 teacher, believed that central administration had imposed the curriculum guide as a goal and that it was valueless for teachers: "I'm positive everybody [in the department PLCs] is not on board with it. But we're doing it because we're told to. It's a whole lot of busy work to prove what we're already doing." Mr. Roses of the Math 10 ELA saw the Math curriculum handbook—again, a more modest demand compared to the ELA document—as a way to show the district's central administration all that the focus PLC and teachers were accomplishing. The Math 10 PLC as a whole turned the compliance work on the document into an opportunity to revise and focus the essential standards, which in turn led to the collective revision of certain lesson activities and formative assessments.

Notwithstanding the variations in PLC and evaluation between and within the case schools, a clear and consistent asymmetry between the values and goals of PLC and evaluation responses had a significant implication when a structural relationship between the two presented an opportunity for punctuation. For teachers could obtain valued PLC benefits and achieve desired PLC goals through the very process of responding to PLC demands. The low value, and in any event incommensurate, products of evaluation were irrelevant to the PLC process. But those PLC responses came in handy when adaptively applied to evaluation.

This explains why Ms. Macallen who opposed Marzano scales was nonetheless grateful that she could simply plug them into her evaluation documents. It is why Ms. Dalmore and Ms. Laphroaig said that the Selective Use and Professional Validation of scales reduced the stress of observation preparation, making it easier and less time-consuming. As Ms. Laphroaig elaborated:

> … [W]hen we say, you're going to come into my classroom, this is the lesson you're going to see; these are things that I'm evaluating. That's not something we are reinventing every year, that we're drawing on work that we've done collectively as a group in order

to make that sort of streamlined.

It is why Mr. Bulliet pointed out that the district's changes to the demonstration of student growth did not increase the burden of evaluation compliance because his PLC was already engaged in such analysis, as well as why many Walker teachers likewise choose PLC goals as evaluation goals. And it is why Ms. Daniels and Mr. Roses were eager in evaluation meetings to promote their work in the Math 10 PLC.

Teachers normally compartmentalized their responses to PLC and evaluation demands as each arose because the goal and corresponding value of responding could be directly, independently achieved. The low-valued evaluation process demanded individual compliance, and only so much as would result in an effective or highly effective summative rating. PLC demands could also involve a compliance goal to some degree, but determining when a group's work product was compliant, and imposing any consequence for non-compliance, were problematic exercises at best. Moreover, beyond mere compliance, teachers' PLC responses often served higher-valued goals like peer collegiality, shared responsibility, and professional learning. It was in this context that teachers would often take opportunities to increase the utility of their PLC responses by decreasing the cost of their evaluation responses. At the same time, the achievement of PLC goals, even compliance goals, were not hindered but could be enhanced, and the achievement of the evaluation goal was made more efficient. Teachers sought, in short, reconcile their agentic responses with structural demands by engaging in punctuated compartmentalization.

CHAPTER 6
DESIGN AND GOAL ORIENTATIONS UNDERLING EVALUATION/PLC STRUCTURAL RELATIONSHIPS SHAPE TEACHER ADAPTIVE RESPONSES

Chapter 6 applies the structure branch of this study's structure-agency framework as it answers my second research question: How, if at all, does the relationship between structures of performance evaluation and professional learning community shape teachers' adaptive responses?

The answer begins in the first section of this chapter, which descriptively compares how each case school and its district differed in their approach to relating PLC and evaluation structures as they were implemented during the school year. Because the approaches were alternative ways of addressing the coherence challenge that evaluation and PLC presented to each administration, the section also explores the alternative design and goal orientations underlying each administration's structural relationship. Based on its means to and objective for attaining coherence, the Woodford administration implemented structures that were by design primarily control oriented with an external goal orientation, while the design of Walker's structures was primarily commitment oriented with an internal goal orientation. The second section examines how the alternative, administratively determined structures and relationships affected teachers' opportunities for punctuation, which involved the interaction and adaptation of compartmentalized PLC/evaluation responses, as well as teachers' use of particular adaptive response strategies to take advantage of those opportunities.

Relationships Between PLC and Evaluation Structures Rest on Administrative Design and Goal Orientations

Chapter 5 showed how and why Woodford and Walker teachers use four kinds of adaptive response strategies when the structural relationships between PLC and evaluation opened opportunities for agentic action. Yet the structural relationships between PLC and

evaluation at the case schools were hardly the same. Further, it would seem that Woodford and Walker had faced the same sort of coherence challenge as they routinized the implementation of a state-mandated evaluation system along with a locally-adopted PLC model. Yet, again, the relationships between structures today differed. Before one can understand how varying structural relationships may affect teachers' adaptive responses, one must know how each case reached its current state.

Woodford "has always had PLCs," Ms. Lauder thought, and its teacher performance evaluation dated to the state law mandate. The district pursued both along parallel tracks until the state law was amended to require that districts choose an evaluation tool from a state-approved list. Woodford chose the "Marzano Focused Teacher Evaluation Model." It was this decision that precipitated a serious coherence challenge between evaluation and PLCs. The challenge was accentuated by the then-existing structure of PLCs and a strong aversion to the Marzano framework among teachers.

District leaders had determined that to improve teaching and learning district-wide, they would adopt a consistent instructional framework providing for standards-based units with defined student outcomes and aligned assessments. Once state law required a choice, and the Marzano tool seemed the better option in the circumstance, "what we said from this [central] office is we wanted to make the best of a bad thing." Ms. Lauder explained further:

> Well, first and foremost, it [evaluation] was implemented because it's state law. … However, rather than it just feeling like a hoop, a compliance hoop, we've really attempted to embed this [Marzano] model into our professional development so that it's seen as a growth tool.

At another point, Ms. Lauder returned to her view of the relationship between evaluation, the district's goal, and teacher reaction to the district's approach:

> We knew we had to comply. In looking at the Marzano elements, they themselves are not inherently evil. And we have to assign a number, and I think assigning the number has been not very helpful in growing teachers. We are hoping that over time the anxiety about those numbers will decrease as people realize that it's not—they're not losing their jobs; it's not changing the way that they do their work—that their anxiety about that will go away, and we'll be able to benefit from having a consistent instructional framework, which I would have done anyway, whether Marzano was here or not.

The bottom line remained, Ms. Lauder emphasized, that "I do want there to be a consistent framework district-wide, and that's the framework we use."

Beginning in the 2017-2018 school year, the district rolled out a three-year professional development plan to advance the overarching goal—a consistent instructional framework district-wide. The coherence challenge would be addressed through the process of implementing the Marzano framework/model in multiple professional development settings and activities. Most significantly, it constituted the required tool in teacher evaluation, and it would be integrated into, on some issues it would drive, the work of PLCs. Plan implementation depended on certain changes in the then-existing systems of evaluation and PLCs, such as the following:

- The district not only trained administrator-evaluators on the use of the Marzano tool, but it also directed them to rate only those elements they actually observed; for the final summative rating, evaluators had to observe and rate all twenty-three elements. Discussions during the pre- and post-observation conferences, as well as other required teacher responses (e.g., self-assessment, professional development plan, lesson planning) were to involve the Marzano tool/model, at least in part.
- For their observations, teachers had to submit a scale (or relevant portion of one) with their lesson plan; the scale would also be a subject during pre- and post-observation conferences. Dr. Sark, for example, stressed that teachers were

> expected, through the evaluation, … that they have a scale and that they have explicit ways to show, tell students where they are in their learning, where they're going, what they're learning, using that scale to talk about that in the classroom. … If you don't have that, you're completely dead in the water on the first, on the planning and the standards based [elements] on the evaluation tool. So basically, that's an explicit connection that we make often, too, with our staff that you have to have a scale, you have to know how to do a scale.

- The district implemented a district-wide calendar of for all professional learning meetings and activities, including those involving content time and collaborative time PLCs, as well as whole building initiatives, cross-level teacher meetings, and district professional learning days. PLC.
- For content or department time PLCs, like the ELA 11 PLC, department heads like Ms. Dalmore, met with building and central administrators before the start of each school year. District administrators at this meeting first gave the department heads feedback on their department's PLC work from the prior year (which had been reviewed over the summer). Ms. Lauder then directly assigned grade-level/content PLC work and required PLC products for completion during the meetings scheduled for the year. Ms. Dalmore described the first such meeting pursuant to the plan (2017-2018), as well as the meeting going into the 2019-2020 school year:

> So the department chairs were instructed at the beginning of this school year, I mean, really two years prior that we needed to start working on aligning the scales of, creating scales that showed proficiency of the skills within … the unit. … So luckily, [Ms. Lauder] is nice and flexible about that [organizing by unit], which is good. But we knew we had to do this.
>
> * * *
>
> … [A]nd they kind of wanted to see where everybody was in the process, so we've been adding our completed or, you know, rough draft I guess, scales to a SharePoint folder that the curricu-

> lum office has access to. So [Ms. Lauder] and [Elementary Curriculum Director] looked over those this summer and gave some, I'd say some descriptive feedback or questioning feedback on the scales that they gave back to department chairs at the beginning of the school year this year, and then gave us like a kind of a checklist of other additional things to accomplish where, you know, like here's where you are, here's where we need to go.

Ms. Dalmore would make a calendar, which she shared with her department's PLCs, that scheduled the goal for each meeting during the year so that the districts expectations for PLC work product would be achieved.

- Because the district was required to provide teaches with professional development on the evaluation tool, the district reorganized collaborative time PLCs (those not under study) to match the Marzano framework instructional domains. Teachers selected one of these cross-grade, cross-discipline groups to work on the designated instructional element over the year.
- District administrators were aware that, as Ms. Lauder put it, "if you talk about the Marzano model, our teachers will scrunch up their faces and say, 'ee-eu,' like they smell something awful." Further, they saw one of the obstacles to implementation as PLC "time on task and feeling ownership over the work. So anytime you centralize work, people feel like it's being imposed on them, even if it's their own work." Consequently, the district did not communicate in Marzano terms, but instead tried to send the message that "we're not doing it because it's Marzano; we're doing it because it's best practice, but it happens to align with our tool as well."

For the 2019-2020 school year, the ELA PLCs were to continue working on the district project of implementing the Marzano instructional framework through their curricular documents, starting with the scales the PLCs had produced the previous year. The ELA 11 PLC

in particular was to revise the wording of its scales "to use those Marzano words specifically as opposed to our own thinking" (Ms. Dalmore) and to review the scales' alignment with the relevant curricular units and the semester examinations. The updated scales were to serve as the foundation on which the PLC was to construct an ELA 11 "course pack." In addition to the scales, the course pack would consist of a narrative description of the course, its units, and its scope and sequence (akin to a standardized master syllabus). Later, common unit assessments and corresponding student work exemplars would be included.

The origin of Walker's PLC/evaluation coherence challenge lay in a precipitating event unlike Woodford's. Walker had adopted the Danielson framework/tool early on in the its implementation of the state-mandated evaluation system. It did not have PLCs until years later when Mr. Creek almost single-handedly brought them to the district. After researching, training and advocating for the DuFour model, Mr. Creek took the opportunity to implement the DuFour PLC model about three years after becoming Walker High School principal. He was joined in the effort by the new assistant high school principal, Mr. Craig, who also became a DuFour disciple. Because Mr. Creek and Mr. Craig were convinced that the DuFour PLC model was the best route to teacher and school improvement, their main aim for the high school its implementation with fidelity, which required a deliberate professional development process to get teacher buy-in and successful experience with the model.

Because building administrators still had conduct teacher evaluations in compliance with state law and district policy, a coherence challenge ensued at Walker High School. Due to its unreasonable demands on administrators' limited capacities, Mr. Creek and Mr. Craig saw evaluation as a matter of compliance even before DuFour PLCs arrived. Worse, to Mr. Creek, evaluation and PLC were "pretty much in direct antithesis of each other, when you're trying to

foster teamwork, and they you're saying you're going to have an individual score … [I]t just pits teachers against each other that are supposed to be working together." Mr. Craig had the same perspective:

> [T]hey don't really mesh together very well. … One is compliance, and ranking and sorting, and the other is collaboration and growth. I really, I mean that's honestly truthful. I mean, one is literally slapping a number on them, and telling them you are a ninety-four point one, and the other is this process of collaboration and growth.

Mr. Craig further explained that for year Mr. Creek and he had been trying to address the coherence challenge, what he called "this pull and struggle between the PLC process and formal evaluation."

District central administration eventually decided to implement PLCs in all its schools. Fortunately for him, Mr. Creek retained almost unfettered latitude in implementation at the high school. The district supported the efforts of Mr. Creek and Mr. Craig through funding DuFour training and approving school schedule changes, but it was and remained "not hands on," as Mr. Creek said, or "hands off," as the assistant superintendent, Mr. Mark, put it. Historically, the district also allowed building administrators a mostly free hand in the conduct of teacher evaluations were conducted. (The district may have started to change its approach in 2019-2020, as noted below.)

Given their view that evaluation was a best an impediment to effective PLCs and at worst a threat, Mr. Creek reasoned as follows:

> So obviously part of it is state mandated that we have to follow the law of the land, if you will. But we don't feel it to be an effective way to really make people into being better instructors. So, you know, I mean, we do pretty much what the minimal law requires us to do to evaluate and still meet the standards that the state is looking for.

(Mr. Creek and Mr. Craig separately told the same story about DuFour taking the same position during a presentation they had attended.) Mr. Creek and Mr. Craig therefore decided to "de-emphasize" evaluation but nonetheless use some of its components "to leverage the things we already want to do." They accordingly implemented the DuFour PLC model through actions such as the following:

- Beyond their own extensive DuFour training, Mr. Creek and Mr. Craig sponsored cohorts of Walker teachers to attend the annual DuFour PLC "institutes." By 2019-2020 over 60% of the teaching staff had gone. They also personally delivered extensive in-school professional development on the DuFour PLC mode.
- Practically every Wednesday, the regular class schedule was modified to create an extra class period, which was used on alternate weeks for PLC and the other main school initiative, a Response to Intervention (RTI) program. Later, mandatory after school staff meeting time was converted to PLC time. A portion of district professional development day time was turned over to PLCs as well.
- PLCs were organized by department first, and then by grade-level/content groups. Some time was allotted for departmental issues, but the smaller PLCs had substantial autonomy to develop their own norms, agendas, processes, etc. There were general expectations and work product requirements. For core subjects like ELA and Math, these included the establishment of course "essential standards," curricular scope and sequence, and common summative and formative assessments, all aligned with the MI CCSS and the SAT. With these in place, per the DuFour model, PLC work was supposed to emphasize student data collection, analysis and comparison.

- Building administration reduced the paperwork and teacher submissions required for evaluation.
- Building administration encouraged teachers to adopt their PLC's goal as their own for evaluation. Administrators in evaluation meetings almost always asked about teachers' PLC work and needs for support. Otherwise, evaluation meetings were short, informal, low pressure, and not overly concerned with individual teachers' performances. It was widely believed that virtually every teacher was at least rated "effective."
- Mr. Creek was aware that his teaching staff took their cues from him, so he and Mr. Craig endeavored to send a strong message about the relative importance of and expectations around evaluation and PLC. As Mr. Creek explained:

> … [W]e've tried to minimize how important that [evaluation] is to us, to our staff. And I think they understand that it's not really what we're looking for, but we have to do it. And, you know, there's certain things we've got to comply with, and we'll do those things. We're not trying to hide anything. We're not trying to avoid anything. But, you know, I think for all, you know, I would use minimally, as we're required by law to do, and we'll do that, but we're not going to focus on it.

For 2019-2020 school year, Mr. Creek created a "PLC Guiding Coalition" of department representatives. They met regularly to address PLC issues and to act as liaisons between building administration and the departmental PLCs. Individual PLCs were to continue working on their self-adopted goals consistent with the administration's general expectations. The district central administration did add several twists to the year, however.

In September, central office issued a revised "Evaluation Handbook," along with additional evaluation and PLC requirements. Administrators were then required to electronically submit a form documenting every informal observation. To tie individual student performance in

each class to the classroom teacher the handbook provided a new classroom standard to the rating categories. PLCs were tasked with producing a curriculum handbook, and in the case of the ELA department, to complete the more extensive curriculum guide, together with a generic student growth assessment. Finally, building administrators were to conduct "mid-year review" meetings with each of their teachers. (Mr. Creek responded to this last evaluation demand with the teacher questionnaire described in Chapter 5.) The ELA department, at least, saw these changes as unwarranted interference with the school-level PLC work.

A comparison of *how* and *why* Woodford and Walker related PLC/evaluation structures as a designed means to a particular end goal reflects two opposing orientations. Woodford structured its PLC/evaluation demands with a *control* design orientation, while Walker employed a *commitment* design orientation (Rowan, 1990; see Ingersoll, 2003; Talbert, 2010). Further, Woodford's overarching goal—one district-wide instructional framework—was outside of or *external* to the related structures of evaluation and PLC; Walker sought to successfully establish the authentic DuFour PLC model in the high school, a goal integral or *internal* to the related structures. The means-end pairing of control-external and commitment-internal orientations reflects an obvious logic, at least in the context of the two cases here. At Walker, PLCs are almost by definition embody the commitment orientation, the DuFour model at the school level particularly so (DuFour, 2004; DuFour and Eaker, 1998; DuFour and Fullan, 2013; DuFour, et al., 2004). The goal is directed internally—i.e., to and for the school whose members are empowered by it. Conversely, a school, still less a district, administration could implement DuFour PLCs by control design, since authentic collaboration cannot be conjured by fiat, not the least because teachers will resist working toward the intended goal (Talbert, 2010; see Hargreaves, 2007). The Woodside central administration, in contrast, adopted a goal that

explicitly involved implementing the same instructional framework for all schools and teachers in the district. Centralized control of implementation is necessary to ensure that all standardize their instructional framework with strict fidelity. Otherwise, the benefit for all units intended by the center cannot be achieved. If a commitment orientation were adopted instead, the inevitable fragmentation would lead to uneven, or maybe no, implementation, likely wasting resources without any significant, scalable or sustainable improvement.

To summarize thus far: Within the last few years Woodford and Walker independently faced a coherence challenge precipitated by a policy determination about a particular goal around teacher and school improvement. The administration in each locale designed its own path to address the coherence challenge and to achieve its policy goal. Both plans required the reconfiguration of PLC and evaluation structures in relationship to each other as an important means to the larger ends. The unsurprising result was different structural relationships based on different design orientations that are directed to different goal orientations.

Each linked pair of orientations—control-external in Woodford, commitment-internal in Walker—conditioned teacher responses to punctuation opportunities, including the use of adaptive response strategies. The next section examines how.

Administrative Design and Goal Orientations Affect Teacher Punctuation Opportunities and Related Use of Adaptive Response Strategies

The first section of this chapter explicated the characteristics of the PLC/evaluation structural relationships, resting as each does on linked orientations, at Woodford and Walker respectively. Table 6.1 comparatively summaries these characteristics.

Table 6.1 Comparative Characteristics of PLC/Evaluation Structural Relationships

District/ School	**Woodford**	**Walker**
Grade-Level/Content PLCs	ELA 11	ELA 10 Math 10

Table 6.1 (cont'd)		
Primary Level Determining Structural Relationship	District Central Administration	School Administration
Primary Goal of Structural Relationship	Implement Marzano instructional framework K-12	Implement DuFour PLC model in high school
Primary Goal Orientation of Structural Relationship	External to school level evaluation/PLC	Internal to school level evaluation/PLC
Primary Design Orientation of Structural Relationship	Control	Commitment
Relative Degree of Teacher Agency Within Structural Relationship	Lower	ELA 10 – Moderate Math 10 – Higher
Administrative Framing/Messaging of Structural Relationship	Integrate common elements of Marzano framework into evaluation and PLC demands to incentivize desired responses	Emphasize priority of PLC over evaluation; leverage evaluation demands to encourage PLC development
Structural Relationship Interaction Opportunities	• Initial online evaluation prompts/ planning • Pre- and post-observation conferences; year-end conference • Standard, scale for observation • Some common assessments for student data collection	• Individual/PLC goal-setting • Pre- and post-observation conferences; year-end conference • Standard, lesson, materials, formative assessment for observation • Common assessments for student data collection • Mid-year review questions and meeting • Individual/PLC common standard, lesson, assessment for and analysis of student data
Other Contextual Factors Affecting Structural Relationship	All teachers belong to two kinds of PLC: departmental-content/grade-level and cross-disciplinary collaboration; activities/tasks of both concern implementation of Marzano framework	New central administration requirements for evaluation and PLCs: (a) "Evaluation Handbook," including new informal observation requirements, standard for rating student growth; (b) PLCs create departmental coursebook; (c) mid-year review requirement; (d) ELA department to create 9-12 curriculum document and generic assessment for student growth

It is apparent that alternative design and goal orientations afford punctuation opportunities—the locks of the Chapter 5 metaphor—that vary in number and nature. In light of

the typology's two dimensions, some prospective keys—the adaptive response strategies of Chapter 5—may fit some Woodford or Walker locks better than others. But before addressing this possibility, there is an overarching consideration: teachers' attitudes toward the paired orientations themselves appear to affect teachers' decisions whether and how much to act in response.

The focus teachers in Woodford's ELA 11 PLC exhibited a range of attitudinal reactions to the district's control-external orientations. First of all, Ms. Dalmore's perspective stemmed from her role as the ELA department head, with its attendant responsibilities "up" to administrators and "down" to her teaching peers. She wanted to fulfill her obligations for both; and despite some ambivalence, Ms. Dalmore felt she should participate in the process on account of the possibility for beneficial results. She explained:

> I think in both cases [evaluation and PLC], especially with this, with the scale and this process that's been introduced about using the scales and creating this working document [course pack] that we're using, I think both that and the evaluation process have required a certain amount of buy in and belief in the process that I am for the most part invested and are willing to, you know, take a risk and trust that there's a purpose behind this that will lead to better teaching.
>
> So I would say in that regard I engage in both with of a willingness to try, you know. And I think that that's, I think is important as being a department chair that I have that kind of attitude, too, like the fact that my attitude will influence how much my department engages in this as well. If I approach it with a really negative or dismissive attitude, that they also will feed off of that. So I think it's more that than anything else.

Holding this position, Ms. Dalmore apparently did not feel compelled to lobby, but she would (and did) "follow the rules" to use the Selective Use and Professional Validation strategies with respect to Marzano scales (including selected standards) and common assessments. Presumably she would have used a Dual Process strategy if an opportunity had existed.

Ms. Laphroaig emphasized her use of the same strategies as were available to Ms. Dalmore. Ms. Laphroaig may have also engaged in Lobbying, though it is unclear. She clearly perceived the Marzano-based connections between evaluation and PLC, and she shared the administration's approach to messaging when interacting with her colleagues. (It was suggested that Ms. Laphroaig and Ms. Lauder were personal friends.) Whether or not she supported the administration's goal, Ms. Laphroaig accepted the control orientation because she was getting along by going along, both with her colleagues and with administration. (She would soon achieve tenured status, and she wanted to maintain her highly effective rating.) Ms. Laphroaig was at the same time seeking efficiency in compliance and wanted to get something professionally positive out of all required tasks.

Finally, Ms. Macallen, a veteran teacher, who thought the pre-Marzano systems of evaluation and PLC were beneficial, reacted to evaluation and PLC work in strongly negative terms. She found PLC tasks like the Marzano scales, which had been "pretty much shoved upon us by upper administration," and evaluation process professionally "insulting," as well as "useless" and "worthless." Ms. Macallen did not see any connections to a larger goal, and she tried to minimize her engagement. Not surprisingly, then, Ms. Macallen had no use for Lobbying or Collective Validation; and since Dual Process was unavailable, she did only that explicitly required—Selective Use in evaluation of the Marzano scales.

At Walker, focus PLCs embraced the commitment-internal goal orientations as their own—at least to the extent of their understandings of DuFour PLCs and authentic PLC work. The focus teachers absorbed the administration's messaging about relationship between PLCs and evaluations. The ELA 10 PLC focus teachers all confirmed that the Walker building administration strongly promoted PLCs but strongly disfavored evaluation. The close alignment

of administration's structural orientations and teachers' attitudes reinforced their ownership of the internal goal and buy-in to the commitment design. The Math 10 PLC focus teachers were in agreement. Mr. Bulliet, a second-year teacher, gave perhaps the best summation:

> And I think that we, here at [Walker], we strike a really good balance of putting less weight on the actual evaluations and more weight on how you grow together in your PLC to help improve your practice. And so I think that's been like the biggest thing that I've seen and the biggest way I think that it's, it's helpful. It's just, that focusing more on the growing professionally through the PLCs and less on, like high stakes evaluations and making them less high stakes, less like, make or break for the evaluations, and more about the working together in the PLCs and trying to grow professionally.

Teachers' did not rest their positive attitudes merely on administrators' statements. Ms. Daniels, for example, noted that the building administration helped teacher prioritize PLC work by "carving out time in our schedule" for it. And as Mr. Roses explained:

> But, you know, based on the effort and the resources poured into the PLC process, it's kind of communicated directly or indirectly that, you know, this is really important, you know, teachers working together and looking at data, and collaborating, and coming up with best practices and the best assessments is way more important than you trying to prove you are a good teacher.

The consensus view did not prevent differences among focus teacher use of adaptive response strategies. These seemed to turn on two main factors, which were discussed in Chapter 5. First, the focus PLCs diverged in what each considered important PLC work and so the state of that work. For example, both would selectively use PLC-determined essential standards and common assessments for the student growth component of evaluation. But only the Math 10 PLC was prepared to use the Dual Process strategy for collecting and analyzing student data from the same common assessment of the same standards at about the same time. The two focus PLCs likewise took different approaches to responding when district central administration

required the production of the curriculum document—a rather control-external sort of intervention. (In fairness, the demand on the ELA department exceeded what the Math Department was supposed to accomplish.) The second, somewhat related factor concerned individual and group level preferences given the commitment design orientation. Not all focus teachers, for instance, employed the Dual Process strategy for their PLC/evaluation goal. Again, in the case of Ms. Trace and the ELA department PLC, an addition factor was the setting, and expanding, of the PLC goal by central administration. There were apparently divergences in the use of the Professional Validation and Lobbying strategies, as well, whether in usual evaluation meetings or in the new mid-year review meeting. In short, perhaps due to the relative inexperience of it focus teacher-members, the Math 10 PLC apparently united to use all adaptive response strategies, whereas the ELA 10 focus teachers had more individualized agentic responses.

In addition to the overall effect of structural orientations on focus PLC and focus teacher use of adaptive response strategies, the four strategies could more or less fit the PLC/evaluation structural relationships at the case schools depending on the dimensional location of each. Thus, on the specific direct/holistic indirect dimension, the former strategies are demand-dependent. This implies that administrative structures can in effect compel or preclude the Selective Use and Dual Process strategies. Woodford required the selective use of Marzano scales in evaluation, but it provided no opportunity for the Dual Process strategy due to the tasks and products the district assigned to the PLCs. Walker by implication had PLCs make selective use of common standards and aligned assessments; but because of its commitment orientation, PLCs and teachers could opt or not for the application of the Dual Process strategy to the determination of student growth. With respect to the dual processing of PLC/evaluation goals, Walker made the

choice entirely voluntary. The holistic indirect strategies of Professional Validation and Lobbying are always available, at least in principle, but they depended on what the PLC had actually accomplished and how the teacher had contributed to it (if one is going to be honest, anyway). Whatever the design orientation then, holistic indirect use depends more on teacher initiative, with the likely effect assayed a case-by-case basis.

With respect to the individual teacher/collective PLC focus dimension, a control-external orientation would seem better suited to promote the former type of strategies (Dual Process and Lobbying), if the structures admit of a reasonable connection with the goal. Administration might be able to specify both sides of a Dual Process, again depending on the goal and opportunity costs for PLC resources. As for lobbying, if Woodside administrators in evaluation meetings specifically question teachers about their contributions to PLC activities or professional learning around the Marzano framework, some teachers might be incentivized to do both (or to explain why they could not do so). The general problem, of course, lies in superficial compliance, half-heated responses, or other forms of indifference or resistance. It might seem easier to control the PLC focused strategies (Selective Use and Professional Validation), but here there is again the issue of determining PLC product quality and responsibility for same. Further, even a quality product does not eliminate the free rider problem of group work. A commitment-internal orientation depends on teacher and PLC buy-in, regardless of focus. The free rider problem does not go away for the PLC-focused strategies, but there may be higher quality if teacher drive the work product. Similarly, the teacher-focused strategies, there will certainly be variation depending on the individual finding the work and goal to be linked and worthwhile.

Rowan (1990) suggests that the implementation of either design orientation could improve school performance. Here, the approaches of Woodford and Walker seemed reasonably

matched to their respective goals, which rest, after all, on divergent theories about what will improve teacher performance. Each approach seems to have worked, at least in the short run, by influencing, though not determining, teachers' agentic behavior in mitigation of similar coherence challenges. There also seem, however, to be significant variation in the nature and, value of teacher and PLC responses and performance, which, even before the pandemic, held uncertain promise for lasting, coherent, and meaningful change.

CHAPTER7
DISCUSSION

This dissertation examines a specific instance of what Robinson, et al. (2017, p. 3) in apt understatement called the "difficult challenge" posed by multiple, concurrent reforms: "[a]chieving coherence within and across the improvement efforts undertaken by any one school[.]" Moreover, because coherence cannot be achieved once and for all, meeting a coherence challenge occurs through a dynamic, ongoing process (Honig and Hatch, 2004; Louis and Robinson, 2012). Honig and Hatch (2004) emphasize research showing how districts can partner with their schools to "craft coherence" through bridging and buffering strategies. And to be sure, administrators in Woodford and Walker did implement related, if varying and changing, evaluation and PLC structures. Nevertheless, this research shows that classroom teachers acting within related structures can and do participate in their own form of coherence crafting. It thus substantially broadens and deepens our knowledge of how and by whom the coherence challenge might be successfully addressed at the grassroots level.

The overarching significance of this study for our understanding of the coherence challenge brings with it several important implications that bear elaboration. In the first section below, I discuss the conceptual and analytical power that a structure-agency framework brings to research, as here, involving ongoing, routinized implementation of related, but coherently problematic, teacher quality measures. The second section discusses the principal implications of the findings explicated in Chapters 5 and 6. What do those findings say about how and why coherence crafting resulted from teacher adaptive responses within the structural relationships that district and school administrations implemented? Finally, the discussion in the final section concerns a different set of implications, those which pertain to policymakers, practitioners, and researchers going forward.

Administrative Structures, Teacher Agentic Responses, and Policy Coherence

Researchers do not lack for theoretical approaches to the study of policy implementation. The choices run a gamut from principal-agent theories to social network theories to institutional theories to some versions of critical theory (Coburn, 2016). A sense-making framework is today commonly relied upon in the study of the implementation of a wide variety of reforms (see Spillane, Reiser, and Reimer, 2002). Indeed, sense-making and similar approaches may be especially well-suited to the understanding of the initial or implementation of largely new or innovative policies (e.g., Coburn, 2001; Marsh, et al., 2017; Reinhorn, et al., 2017).

This study examines a different sort of implementation in a different sort of context. Woodford and Walker took separate routes to the institutionalization of related routines for the conduct of teacher evaluation and PLCs year after year. The routines at each site differed in design and goal orientation, as well as organizational detail. At both, however, the development of the related implementation routines at each school was shaped by the coherence challenge posed by the concurrent implementation of an externally imposed and an internally adopted teacher quality policy. By the time of this research, well-established structures and processes of evaluation and PLC, while still subject to modification, consisted of certain demands to which teachers responded. Given the institutionalized and iterative nature of the yearly re-set of PLC and evaluation in significant part to manage a coherence challenge, this research applied a structure-agency approach (Coburn, 2016; Rigby, et al., 2016). Under the guidance of the research questions, the chosen framework focused data collection and analysis on teachers' agentic responses to the structural demands of each policy that administrators in each case had established and implemented in relationship to each other.

If "the question of policy implementation is fundamentally about the relationship between social structure and agency[,]" the results of this study demonstrate the potential utility and versatility of a structure-agency framework for research like that here involving comparative, multi-level coherence challenges arising from interacting responses and structures. The strength of the framework is reflected in the findings about each component of "punctuated compartmentalization," but it is foregrounded in the findings of how punctuation works. As the lock and key metaphor for punctuation in Chapters 5 and 6 suggests, the structure-agency framework facilitates the analysis and understanding of each of punctuation's categorical elements—i.e., policy demands and policy target responses—as well as how they fit together or interact to produce (or not) the policy outcomes of interest. Furthermore, structure-agency approaches can elucidate the influences of structures and actors across multiple organizational levels (Rigby, et al., 2016; see, e.g., Woulfin, 2016). The coherence challenge Woodford and Walker faced arose in large measure due to how the implementation of PLC and evaluation—their structural demands, agentic responses, and interactions—spanned multiple levels within the case schools (school, PLC, teacher) and between the case schools and the larger education policy environment (state, district, school).

These structure-agency advantages stand in stark contrast to the application of frameworks that, whatever the terminology, focus on structure or agency to the exclusion of the other. Because a focus on one or the other often "results in an emphasis on one organizational level," such an alternative approach "often ignores other potential conditions and interactions that may be instrumental in how the policy plays out" (Rigby, et al., 2016, p. 296). Here, for example, a structure-only focus could miss or misconstrue how teacher PLC/evaluation responses interact through the use of adaptive response strategies. A focus only on teacher

agency in applying adaptive response strategies could overlook or underestimate how different structural orientations influenced the adaptation of responses, when adaptation occurred, and by whom. It is in the integration of structure and agency that is peculiarly illuminating about policy implementation in educational settings (Rigby, et al., 2016).

The tools of a structure-agency framework "provide a way to uncover the microprocesses by which social structure influences action, and how action, in turn, influences institutionalization" (Coburn, 2016, p. 473; see, e.g., Russell and Bray, 2016). The instant research utilizes this feature in its comparisons of routinized PLC and evaluation between cases schools and among the embedded PLCs with their teacher-members. Comparison in this context requires the analysis along meaningful dimensions of varying administrative structures and corresponding teacher agentic responses. Chapter 5 proposes a typology of teacher microprocesses denominated as four adaptive response strategies organized within two cross-cutting dimensions. This result emerged from a structure-agency analysis that compared focus PLC and focus teacher agentic responses to the punctuation opportunities opened by structurally related demands. Similarly, Chapter 6 finds that a comparison of the opposing Woodside and Walker structural design and goal orientations helps explain how the resulting structural relationships based on the control-external versus the commitment-internal pairings may influence teachers' use of adaptive responses.

Finally, a structure-agency lens benefited the examinations of how Woodford and Walker administrators and teachers attempted to address the PLC/evaluation coherence challenge in the local contexts of Woodford and Walker. Efforts to craft coherence almost by definition involve the recursive interaction of policy structures and agentic responses (Honig and Hatch, 2004). A structure-agency focus highlights the recursive nature of these interactions, which in Woodford

and Walker have effects on administrator and teacher responses to the structural demands those in each position confront. For example, the more teachers in Walker opt for the Dual Process strategy when choosing their evaluation and PLC goals, or the teacher-members of the Math 10 PLC use Lobbying to garner positive recognition of its work, the more those structural relationships are reinforced, as a component of evaluation becomes more about PLC work and results. On the other hand, when the Walker district directed mid-year reviews and a new standard for determining teachers' student growth rating, Mr. Creek responded to what were from his position structural changes. His actions—namely, making the teacher questionnaire the basis for discussion and encouraging a Dual Process strategy for analyzing student data—in effect opened new punctuation opportunities that, if taken by teachers, may change PLC and/or evaluation structures.

Coburn (2016, p. 473) argues that "[m]ost policies in education seek to foster change in organizations (schools and school systems) by shaping and channeling individual behavior in particular directions." That is certainly the case at both the individual and embedded group (PLC) levels in the routinized, related functioning of evaluation and PLCs at Woodside and Walker. Coburn (2016, p. 473) continues:

> Theories of structure and agency provide conceptual tools for understanding when and how aspects of the social structure (in this case policy) influence individual action and interaction and how that, in turn, reinforces or alters rules, roles, and relationships in schools and school systems.

Yet researchers at present seem to underutilize a structure-agency approach. Schooling during the unprecedented consequences of the covid-19 virus entails radical disruptions in routinized structures, responses and interactions at all levels in all processes, resulting in a cascade of coherence challenges, many of which may persist even as the pandemic abates. Especially in the

contemporary context, then, this study suggests that it might prove fruitful to exploit the strengths of a structure-agency approach.

Coherence Crafting Through Teacher Adaptive Response Strategies Within Related Structural Demands

Scholars have long recognized the signal role played by "street-level bureaucrats," who in effect make policy through the very process of implementing it (Lipsky, 2010 [1980]; Weatherley and Lipsky, 1977). The phenomenon occurs because "the work of street-level bureaucrats is inherently discretionary"; their work is subject to the policy structures that specify what they are to implement, but as a practical matter, "street-level bureaucrats are constrained but not directed in their work" (Weatherley and Lipsky, 1977, p. 172). In the education policy field, the role often falls to teachers and administrators at the school level because it is there that they must exercise constrained agency to implement policies designed, adopted, and promulgated by the system's higher levels (Knapp, et al., 1998; see, e.g., Donaldson and Woulfin, 2018; Russell and Bray, 2013).

This multiple case study does not involve how teachers as school-embedded street-level bureaucrats responded to the demands of *one* policy; rather, it seeks to explain how teachers responded to a coherence challenge—the concurrent, iterative implementation of *two* very elaborate, very different teacher quality improvement policies. This study broadens our understanding from teachers' agentic actions altering a policy in implementation to their agentic actions contributing to coherence between policies as implemented. It extends and augments our knowledge of how an "old" phenomenon in implementation works in contexts involving the implementation of multiple teacher quality reforms. At the same time, given the contexts of the case schools, this research both confirms and revises aspects of how "coherence crafting" itself is conceptualized.

Honig and Hatch (2004) identify the coherence challenge many districts and schools face during the current multiple reform wave. Their literature review identifies two categories of response, which they refer to as "outside in" and "inside out." The former concerned "first-generation systemic and standards-based reforms [that] have addressed the challenge from the point of policy origin," including district central offices; the latter relied upon "second-generation systemic reforms [that] focused on solutions within schools." Since they find both approaches wanting, these researchers instead argue for a combination approach. Based on the lessons they draw from the literature, Honig and Hatch (2004, p. 19) define coherence as a

> process of negotiation whereby school leaders and central office administrators continually craft the fit between external policy demands and schools' own goals and strategies and use external demands strategically to inform and enable implementation of those goals and strategies.

They then discuss and give a research basis for administrators to engage in bridging and buffering activities aimed at coherence crafting. In short, Honig and Hatch (2004) conceive of coherence crafting as occurring between the district and school levels, and as mainly consisting of administrator responses to external demands in relation to local goals. Some research since has focused agentic responses to higher level structural demands by actors at the school level, such as principals (e.g., Donaldson and Woulfin, 2018) or teachers (e.g., Robinson, 2012) or both (e.g., Chingara and Heystek, 2019; Stosich, 2018).

The administrators responsible for the two case schools under study were ultimately trying to improve teaching and learning when they precipitated and addressed a coherence challenge arising from the regular, school year long implementation of an imposed system of teacher evaluation and a chosen model of PLCs. Both appreciated the problematic nature and likely effects of the external policy mandate, but both also saw an opportunity to structure the

evaluation/PLC relationship in implementation to advance a favored goal, each a different intermediate means to student achievement and school improvement. The alternative pairs of orientations themselves rested on divergent convictions about the way to change teacher practice, thereby improving teacher quality. Indeed, Indeed, Woodford's control-external orientations seem akin to the first-generation response to a coherence challenge that Honig and Hatch (2004) identify, while Walker's commitment-internal response seems a nod to their description of the second-generation response. In any event, Woodside and Walker administrators did attempt the kind of coherence crafting between levels that Honig and Hatch (2004) advocate, even though they utilized different bridging and buffering strategies.

To be sure, Chapter 6 substantiates that administrator coherence crafting efforts consisting of interacting PLC and evaluation demands did influence teacher and PLC responses, with variation due to different structural orientations. But however necessary, such one-sided efforts, being administrative and structural only, were insufficient. They could not alone "inform and enable the implementation" of the local goals in a coherent manner. To the contrary, Chapter 5 shows that teacher agentic responses played an indispensable role, made an essential contribution, to the crafting of coherence.

Several significant implications follow from the extension of "coherence crafting" from administrators to teachers (and teacher groups like PLCs). First, if both administrators and teachers participate in coherence crafting, it necessarily depends on two categories of actor that vary in structural context and degree of agency. Successful coherence crafting, at least as seen through a structure-agency lens, must therefore attend to two sets of interacting structures and agentic responses—those of administrators and those of teachers. The sets themselves are

recursively inter-related since the results of the administrator interactions become the structures in the teacher interactions.

Second, from their differing positions, teachers and administrators will not necessarily share a perspective either on what counts as "coherence" or what it is for. As Chapter 6 discusses, administrator demands reflect their design and goal orientations, but as Chapter 5 demonstrates, teachers act based on how they see the goals and benefits of responding. Coherence crafting is a doubly dynamic process of trying to sustainably align potentially divergent interests toward a mutually acceptable goal. Even if alignment occurs and coherence is achieved, they may be only at a superficial level or only for the short run. In Woodford, for example, the district was able to compel coherence in the PLC production and teacher use in evaluation of the Marzano scales, but one wonders how far that advances the district's larger goal. While the Walker Math 10 PLC accepted the Dual Process opportunity regarding its analysis of student data, but many others apparently were not interested, despite its contribution to PLC/evaluation coherence.

Third and last, a broader, more complex, actionable understanding of who must participate how to meet a coherence challenge may be more robust, it remains at least as susceptible in practice to shocks or other changes in organizational environment or relationships. An obvious example with respect to routinized PLCs and evaluation would be if substantial lay-offs were required. Less extreme but still problematic in Walker's context would be continued evaluation process or PLC work mandates from the central administration. Obviously, the pandemic has had a catastrophic effect on the implementation of all manner of reforms and whatever coherence challenges they engendered. Nonetheless, even in such circumstances, this study's shows that, if it is to have a realistic chance, coherence crafting requires the participation

of teachers and must account for their responses, while administrators and school leaders must with intentionality provide for the positive interaction or integration of policy demands.

This research develops a more sophisticated and nuanced understanding of how teachers, individually in evaluation and collectively in PLCs, respond to the institutionalization of related policies, the demands of which have converged in the case schools. But further, the interacting demands in design and goal orientation, together with the adaptive responses in motivation and effect, can both address a coherence challenge. This study thus illuminates key features of a common coherence challenge, yields analytical insights into alternative efforts to mitigate the challenge in context, and points to ways for assessing how recursive demands and responses might converge toward or diverge from policy coherence in routinized implementation over time.

Implications for Policymakers, Practitioners, and Researchers

It seems inevitable that education policy prescriptions are altered in their implementation. In the agentic hands of street-level bureaucrats, even seemingly "simple," straightforward policies can be transmogrified into something quite different from what was written, let alone intended. Like parents going from child to children, the challenge of implementing one reform with a balance of fidelity and flexibility is non-linearly compounded by the implementation of a second scheme meant to address the same problem through different means. In this study, the case schools' routinized implementation of evaluation and PLC together engendered an ongoing coherence challenge in addition to the "ordinary" challenge of implementing one or the other alone. How the interacting policy demands and responses played out has abstractly similar but practically different implications for policymakers, practitioners, and researchers going forward.

Policymakers and local leaders contribute to policy churn (Hess, 1999). While that may be politically irresistible, they should when designing and initially implementing new policies try to anticipate and provide for demand relationships and interactions with other policies in a relevant range of existing local contexts. They should consider explicit and implicit dimensions of policies that may in principal or practice conflict or otherwise create coherence challenges. For example, the state legislation mandating the evaluation system to which Walker and Woodford were subject on its face stated basically incompatible formative and summative goals. In most districts, including Walker and Woodford, the formative goal was virtually impossible to implement from the beginning, which left the summative goal of individual accountability in direct conflict with the collective accountability and collaborative goals of PLCs. State policymakers could help mitigate potential incoherence were they to rebalance the formative with the summative components and/or give districts more discretion in emphasizing evaluation for formative purposes—for instance, allowing districts to choose from a greater range of evaluation tools or to modify the standard ones, or reducing the time-consuming requirements for administrators. At the very least, policymakers should avoid mandating evaluation components such as VAMs that likely exacerbate coherence challenges (Johnson, 2015).

When a district and school practitioner leaders also act as policymakers—for example, when Woodford chose the Marzano tool and made its professional development plan—they might tailor their design and goal orientations, as well as their messaging, to anticipate teacher responses and to mitigate the effects of more abrupt changes in existing PLC or evaluation demands to which teachers are accustomed. Somewhat along the lines Mr. Creek was attempting in Walker, state and local policymakers might explicitly provide for the partial or complete integration of PLC and evaluation demands as several scholars have proposed (Jones, et al.,

2016; Woodland and Mazur, 2015; see also Reinhorn, et al., 2017). The extent to which such moves would enhance coherence would depend on, among other things, the local history around reform adoption and implementation, availability of suitable professional learning, sensitivity to reliance interests, and staff and leadership composition and experience.

Regarding policy design and implementation planning, this study suggests that local leaders and policymakers should be mindful that many stakeholder have some degree of agency in implementing and responding to policy demands and that their responses stem from interests, goals and values that do not necessarily match or align with those at other levels. Thus, district and school leaders should expect disparate efforts at coherence crafting at all levels and recognize the trade-offs among a range of demand locks and response keys. Demand and response interactions that may continue or arise after initial implementation and that may affect coherence, must be considered and addressed (Knapp, et al., 1998). Even absent formal integration of the demands, both case schools might review their professional development plans to make their related structures, anticipated responses, and ultimate goals clear and persuasive to teachers and administrators.

Among those contributing to policy implementation under study here, it was of course teachers individually and in their PLCs that were this dissertation's main subjects. Perhaps it is hardly surprising, then, that several implications for teachers may be the most suggestive, even counter-intuitive. This study shows that, when confronted with PLC and evaluation demands, teachers regularly exercise agency in response. Teachers' interaction of responses in the application of adaptive response strategies contributes to the ongoing task of crafting coherence, even when schools may implement structures with opposing orientations that influence teacher

responses. Teacher agency, in short, has a very real effect on the processes and results of stable, routinized implementation.

Yet, one implication of the context, opportunities and ways by which teachers *can and do* play agentic roles is that they *could and should* do more in the routine enactment of related PLC/evaluation to change both for the better going forward (Conley and Enomoto, 2005; Horn and Little, 2010; Feldman and Pentland, 2006). Particularly because all routines entail structural or "ostensive" and agentic or "performative" components (Feldman and Pentland, 2006; Horn and Little, 2010), the implication is that teachers could proactively "lean in" to opportunities for enhancing and extending the recursive interactions among the structural demands and agentic responses of PLC work and evaluation (Coburn, 2016). The intended effect would be to increase PLC/evaluation integration and coherence (Honig and Hatch, 2004), focusing both on effective professional learning and practice improvement (Reinhorn, et al. 2017; Taylor and Tyler, 2012; see also O'Day, 2002; Reinhorn and Johnson, 2015). Teacher/PLC efforts along such lines could closely align with and reciprocally support administrators' interests in effective, efficient, and coherent routines for teacher improvement. But more important for teachers' work, these efforts could exemplify the benefits in routinized implementation of affording scope for teacher professionalism and giving impetus to professional accountability.

Teachers might Individually and collectively as PLCs, teachers might start by recognizing their existing agency, and reconceptualizing how and why they might use it. In the present context, teacher agency can be a two-edged sword: teachers asserting a certain autonomy in pursuit of their interests at the same time assume a certain responsibility for the consequences. While teachers could opt for the stances of passive vessel, indifferent cog, or hapless victim, they could instead avoid or reject the unproductive, if understandable, rut of compliance-focused,

wholly reactive engagement with routinized demands like those of PLC and evaluation (see, e.g., Knapp, et al., 1998; Lane, 2020; Talbert, 2010). Teachers might then with intention foster PLC/evaluation integration and cohesion through agentic responses such as the following:

- Teachers could reframe assigned PLC tasks or goals to serve or enable authentic PLC purposes. The Walker Math 10 PLC, for example, took the administrative requirement of a curriculum guide as an opportunity to revise the geometry essential standards and related common lesson plans and assessments. The PLCs in the Walker ELA Department treated a similar, albeit more onerous, assignment as an ad hoc task with which to comply. Mr. Roses saw the midyear review as an opportunity to employ the Lobbying adaptive response strategy with the superintendent, whereas Ms. Daniels framed it as an administrative failure of formative feedback. Woodford's Ms. Laphroiag treated the preparation of the assigned ELA 11 narrative as a means to protect classroom teacher flexibility in the individual enactment of the common curriculum, whereas Ms. Macallen saw it as an administratively imposed waste of time.
- Teachers could identify or propose new opportunities to directly integrate authentic PLC work or products with evaluation components or metrics. When the central administration in Walker required a new standard for determining student growth in individual teachers' classes, the Math 10 PLC planned a Dual Process strategy for the collection and analysis of geometry students' data for evaluation and PLC purposes. Teachers' option in Walker to adopt a PLC goal as an individual evaluation goal also involved a Dual Process strategy such that participation in a collaborative process and outcome contributed to one's summative performance rating (see Jones, et al., 2016;

Kimball, 2002). Departments or PLCs could advocate for similar opportunities with respect to the Selective Use strategy—for example, creating a bank of "PLC-approved" formative assessment, which teachers might use during an observation to "automatically" satisfy elements of the evaluation tool. The Danielson evaluation tool includes a "Framework Cluster" of standards for "Professionalism," the Marzano model has a domain comprising several standards for "Professional Responsibilities." Particularly if administration encourages it, teachers in a PLC could legitimately submit evidence of effective collaboration and successful PLC work to support a higher effectiveness rating.

- Teachers in PLCs could engage in professional development (PD) activities that relate to evaluation components. For example, PLCs could undertake book studies, lesson studies, or classroom observations around the Danielson or Marzano frameworks, *not* as evaluation tools, but as clarifying and helpful instructional taxonomies (which was their original purpose). Ms. Daniels and Mr. Bulliet described how their Math 10 PLC arranged to observe the same lesson on the same day in each of the PLC's three members' respective class. Ms. Dalmore in Woodford describe how her collaboration time PLC engaged in this sort of work, but it was a larger, cross-disciplinary group that did not focus on ELA curricula or pedagogy. Ms. Trace recalled benefiting from the Walker PD on the Danielson framework, but it was conducted as basically a "one-shot" whole school and district PD when the district adopted the Danielson tool. Teachers might in addition decide to familiarize themselves with the research underlying such frameworks, or research related to other components of evaluation— student data analysis, for example. Teachers could

identify gaps in their professional learning which PLCs (or PLCs meeting as departments) might address. Mr. Roses, for instance, expressed a strong interest in PD on accurate and effective student data collection and analysis, which would be a highly relevant topic for the Walker Math 10 PLC. Similarly, PLCs could engage in inquiry around self-chosen problems of practice and their relations to the relevant instructional framework.

Obviously, some PLC activities would require the approval and support of administration, which may have a strong interest in doing so. Besides enthusiastic permission for the initiative, Mr. Creek's support for the Math 10 teachers' classroom observations mentioned above consisted of two half day substitute teachers. It was nonetheless the sort of PLC professional learning activity that he did ask the teachers to present on their experience to the rest of the staff, which they were happy to do. Administrations without much additional significant monetary cost might likewise encourage and facilitate some flexibility and differentiation in assigned or expected PLC tasks or products based on specific PLC contexts, needs and goals. PLC-driven initiatives, innovations and experiments might be encouraged and results promoted, at least on a limited or pilot basis. But even on their own, departments and PLCs in many schools, including Walker and Woodside, could reconfigure the time and other resources they already control to facilitate skill and content specific PD. Further, especially with administrative authorization, teachers could include their PLC-based PD in response to the professionalism domain of the evaluation tool.

- Teachers in general, and especially teacher leaders, could prioritize the use of strategies like Lobbying and Collective Validation whenever possible in interactions

with non-PLC colleagues and administrators. Both Walker and Woodford focus teachers had several opportunities during the evaluation process to use both, although only some like Ms. Laphroiag and Mr. Roses did so. Beyond increasing the individual use of these strategies, teachers in PLCs and departments might coordinate the subjects and "pitches" such uses, as for example, when Ms. Forester in her midyear review meeting and related questionnaire advocated for a reduction in PLC time devoted to administrative matters, or the Math Department's desire to switch from the required DuFour videos to another form of PD. An approach like this might be especially effective were it combined with the further integration of PLC/evaluation PD, processes (e.g., student data analysis) and products (e.g., curriculum guides).

Furthermore, there are in addition other forums where teacher leaders could apply the strategies on behalf of teachers and PLCs. For example, at Woodford, there were the annual August meeting of department heads with building and central administrators to determine PLC tasks, products and goals. The Woodford principal, Mr. Regal, also conducted regular meetings of administrators and teacher leaders. Walker's principal, Mr. Creek, started a PLC Guidance Committee, and he had a practice of meeting with specific PLCs and department heads at their request.

Building on existing opportunities and strategies, the list above reflects the kinds of agentic responses teachers in PLCs might devise and execute to their own and their schools' benefit. The general approach might counter the deleterious effect of routine external accountability demands that Wills and Sandholtz (2009) in a different accountability context first identified as "constrained professionalism," and assert, or even promote, teacher professionalism in learning, performance, and accountability (Louis, et al., 1996; Talbert and McLaughlin, 1996).

In schools like Walker and Woodford, teachers have a foot in the door; now they need to purposefully walk through it.

With respect to the implications of this work for researchers, this study begins to fill a gap in the literature, which currently has little to say about the interaction of multiple policy demands and related responses. Its design, analysis and results point to potential future research addressing similar questions in similar contexts. Most closely linked to the instant research, future research might investigate such questions as these: To what extent, if any at all, may the typology proposed here be generalized? Might other typologies of teacher adaptive response strategies, based on other dimensions, better fit or explain the interactions of and responses to convergent policies? Similar research questions might be asked about how this research interprets the nature and effects of structural demand design and goal orientations. For example, which orientations are more effective in the longer run in sustaining the intended teacher responses and achieving the larger goal? Are there other orientations that might work better in particular contexts with particular goals?

More generally, researchers might examine the routinized implementation of other accountability or teacher improvement policies for how variation in the integration of demands compares or relates to variation in teacher response adaptations. Building on this study's application of a structure-agency approach, researchers might uncover other mechanisms or microprocesses by which teachers contribute to the crafting of coherence, whether along with or in opposition to the efforts of administrators or other higher-level leaders. Increased knowledge along such lines might facilitate the investigation of whether and how changes in coherence through teacher participation relate to teacher and school improvement.

Another possibility that perhaps warrants investigation is the presence of other factors that significantly influence how and why individual teachers and PLCs exercise their agency to compartmentalize or punctuate responses that hinder or enhance coherence. In this study, for instance, some evidence hinted that probationary or early career teachers more frequently saw and acted upon punctuation opportunities and adaptive response strategies as compared to more veteran colleagues.

Finally, future scholarly work to broaden and deepen research in this field could also help inform policymakers and practitioners seeking to promote coherence (or mitigate incoherence) in multiple policy implementation. Longitudinal studies might elucidate how and why coherence crafting, a dynamic process after all, evolves over time in multiple organizational levels and among various stakeholder groups. Do better coordinated coherence efforts over multiple iterations contribute to implementation fidelity, policy sustainability, and/or desirable policy outcomes?

Limitations

The limitations of this research are mainly inherent in the qualitative nature of its methodology in design and implementation. To begin, given the study's purposeful sampling strategy, at no level—district, school, PLC or individual—were participants randomly selected; none is necessarily representative of the larger population. Further, since at each level individuals voluntarily agreed to participate, often with administration support, the relatively small sample this study comprehends, may differ in significant ways from another potential group of participants. To the extent possible, I tried to address this limitation by taking those steps discussed above to strengthen this study's external validity. This includes the mutually reinforcing design elements of "how" research questions calling for explanatory answers, a

multiple case study with several embedded units, and the replication logic of the purposeful sampling strategy. I emphasize in particular the proper and fruitful focus on explaining processes emerging from, and seeking analytical generalizations about, varying agentic responses to varying structural demands of evaluation and PLC (Maxwell, 2014; Firestone, 1993).

This study's sampling strategy limits the generalizability of its findings in another way, one stemming from the necessarily delimited context of the case schools. For reasons discussed in Chapter 4, the sampling strategy sought case high schools implementing routinized systems of teacher evaluation and PLCs that were "typical" or "middle of the road" with respect to common contextual characteristics. The strategy avoided both especially high and low performing schools, as well as those subject to high accountability pressure, financial distress, or other special externally determined statuses. Findings based on the resulting sample might have limited generalizability in "non-typical," but important contexts. Contexts characterized by high accountability pressures might be of particular concern because a large proportion of contemporary education reform involves the application of strong accountability policies (e.g., Wood, 2007).

The external validity of case study research is strongest in cases like those in the actual sample. The sampling strategy of Chapter 4 suggests that many other schools are "typical" in the case school sense: school level accountability, teacher evaluation, and PLCs are routinely implemented in the absence of significantly greater than average accountability pressure. Years of data show that teacher evaluation is very rarely a high-stakes threat to the vast majority of teachers. Nor do school level accountability systems seem to produce realistic threats of significant sanction for many schools like the case schools here. Nonetheless, schools like Walker and Woodford must comply with accountability mandate to maintain or increase

measured outcomes, deal with ongoing coherence challenges, and devote resources to teacher and school improvement efforts. Mandates for teacher evaluation and school level accountability have been shown to result in significant effects even in the absence of truly high stakes, whether by design (Taylor and Tyler, 2012; Thiel, Schweigher, and Bellmann, 2017), or by schools' attaining sufficient performance levels (Lane, 2020; Marsh, et al., 2017; Johnson, et al., 2018; Tichnor-Wagner, et al., 2016).

Other schools and teachers are subject to different demand structures, and certainly some are more susceptibilities to external pressures. Cleary there are school contexts where powerful influences operate such that the particular processes and outcomes found at Walker and Woodford would be unlikely to occur. Nonetheless, there is reason to believe that *some* degree of generalization from this study's broader findings may be warranted, even in contexts with high accountability or similar external pressures.

This research examines a general problem for educators across accountability contexts: What happens when divergent policies like PLC and teacher evaluation are implemented together, thereby engendering a coherence challenge? The literature on the strategies employed to address the problem do not depend on presence or absence of any particular external policy demand, but rather on certain processes between a district and its school in the form of external policy demands (Honig and Hatch, 2004; Knapp, et al., 1998; see Louis and Robinson, 2012). Further, while empirical research on the implementation of multiple reforms shows variation in the actions and results of coherence crafting efforts (whether or not using that terminology), researchers have nonetheless found consistent patterns of demand/response interactions by the policy targets that do *not* appear to depend on variation in the degree of accountability pressure at the individual or school levels (Louis and Robinson, 2012; Marsh et al., 2017; Reinhorn, et

al., 2017; Reinhorn and Johnson, 2015; Stosich, 2018). This literature suggests that a context including high accountability pressure does not preclude all external validity for this study's basic findings about the participants (administrators, teachers) and processes (paired design-goal orientations, adaptive response strategies) that can contribute to crafting coherence.

To be sure, if significant accountability pressure (or perceived pressure) were introduced to the case schools' contexts, it might very well affect the specifics of how teachers interact and adapt their responses to related PLC/evaluation demands. External validity may be especially restricted for the more detailed aspects of punctuated compartmentalization or the four adaptive response strategies identified here. Perhaps compartmentalization is highly susceptible to certain threats of accountability sanction, or maybe external pressures adversely affect the opportunities for or content of teacher adaptive strategies like Lobbying or Selective Use. One could imagine external pressure so pervasive and sustained as to overwhelm or crowd out all other agentic considerations, structural relationships and interaction processes, including those identified and explained in this study.

But for such relatively extreme cases, however, the generalizability limitation is more ambiguous, for it depends on the nature and degree of effect that any external pressure might have. For one example, the effect on PLCs of school level accountability might differ significantly from the effect of high individual evaluation pressure. Another example: if accountability pressure rendered true compartmentalization effectively impossible, would that necessarily end teacher punctuation that involved similar timing, interaction processes and adaptive responses—say, a recognizable but modified form of Lobbying or PLC product for Selective Use?

Although these examples ultimately pose empirical questions for future research building on this study, the uncertainty about external accountability effects may be due to the persistence in alternative contexts of more fundamental features of the instant research context. These might include, first, that in the face of an ongoing coherence challenge, teachers subject to related PLC/ evaluation demands have opportunities and agency to interact and adapt their responses within administrator implemented structures. Moreover, facing the same coherence challenge, administrators implement related PLC/evaluation structures as a means to goals larger than accountability compliance. These contextual facts may obtain whether or not external pressure exerts some degree of influence on the case-specific demand structures and teacher responses. Thus, the research findings here that identify and explain the *kinds* of processes involving demand/response interactions, the *kinds* of teacher actions contributing to coherence crafting, and the *kinds* of administrator design and goal orientations may be of more general applicability in some alternative contexts that involve significant accountability pressure.

A second kind of limitation relates to internal validity. Although a great deal of observational data was collected, it is of course finite. The data collected does not capture all the potentially relevant PLC interactions or any of the administrator/teacher interactions involving evaluation. It could be, then, that behavior during relevant but unobserved occasions might differ from what the observed evidence shows. Even the observational data that was collected involves two related limitations. First, the happenstance of which focus teachers attended which PLC meetings may have affected the nature and content of the observations. Second, my presence may have affected focus teacher and PLC conduct due to the well-known "Hawthorne Effect." (At Woodford, the presence of interns may have affected the behavior of focus teachers as well.)

I employed several strategies to mitigate these potential internal validity threats. First of all, I with intention conducted observations during the most likely times the focus PLCs would be meeting separately. I was able to take these best opportunities, spread over more than five months based upon the formal schedule of departmental PLC time combined with focus teachers' information about the departmental agenda and their grade-level/content PLC plans. (Recall that at both case high schools, during formally designated PLC time, the focus PLCs would meet once their departmental business had concluded.) At Woodford, but for a content time PLC meeting before students returned in September, I observed all the scheduled grade-level/content focus PLC meetings during the first semester, as well as the first such meeting of the second semester. Because there were two focus PLCs at Walker, I more-or-less alternated observations, depending on advance knowledge about the scheduled departmental work. I nonetheless observed for a significant number and duration of ELA 10 and Math 10 meetings, especially since both invited me to observe voluntary meetings, or voluntarily extended time for meeting. These invitations too are one indication of the rapport I built with the focus teachers, whose individual and collective conduct was remarkably consistent and persistent over the observations and in other settings (e.g., informal chatting, interviews). I did not observe any indication of unusual or inconsistent behavior, including that referencing or alluding to my presence. To the contrary, observed behavior—and so it seems likely unobserved behavior in the same routinized contexts—was consistent with the other sources I used to triangulate among all the collected data. Finally, observational data plays more supporting, corroborating, and triangulating roles with respect to research findings; no finding depends solely, or even principally, on observational data.

These limitations notwithstanding, this research, through its multiple case study design, diverse data sources, and iterative analytical techniques, gives substantial insight into both interacting elements of its agency-structure framework—i.e., how teacher responses to PLC and evaluation demands adaptively interacted, how those adaptive responses were shaped by the way those structures were related. The well-supported findings discussed in the next two chapters identify and explain non-idiosyncratic agentic strategies and structural characteristics and orientations that may apply in typical school contexts, particularly since institutionalized teacher evaluation systems and PLC models often share interacting components and incentivize like teacher responses. At the least, the results provide a basis for hypothesizing and testing in further research, even in research contexts where other factors, such as substantial accountability pressures, may independently influence outcomes.

Conclusion

This dissertation takes a structure-agency approach to the empirical investigation of on an ongoing coherence challenge stemming from the implementation together of PLCs and evaluation. The convergence of similar ends, differing means, and shared periods of routinized re-enactment results in a complex of inter-relationships among two agentic responses, which can influence each other, and two structures, which can interact with each other. A significant contribution of this dissertation follows from its identification and description of how teachers adapt their otherwise separate responses to the related demands evaluation and PLC.

This contribution leads to another: a broader, more elaborated conception of the participants who engage in and mechanisms constituting "crafting coherence" (Honig and Hatch, 2008) between the two widespread, well-established teacher quality policies. By showing that teachers in PLCs can and do participate in meeting the coherence challenge, this study extends

our understanding of who crafts coherence and how. Teacher coherence crafting involves dynamically adapting responses through the application of distinct strategies developed within the related structures of evaluation and PLC. Thus, teacher agentic responses to multiple teacher quality reforms, regardless of source or intention, will from the bottom up interact with the structural efforts to attain coherence at the school and district levels. Finally, this dissertation completes the structure-agency recursive dynamic by examining how the related structural demands of PLC and evaluation, which rest on particular administrative design and goal orientations, will in turn shape teacher adaptive response strategies. Next year the same routine, but not necessarily the same demand structures and agentic response will play out again, except that …

The covid-19 pandemic struck Michigan shortly after the principal data collection for this dissertation concluded. The pandemic upended virtually everything about schooling, including teacher evaluation and PLCs, the routinized implementation of which was and remains in a sort of policy limbo. Nevertheless, these policies are too embedded to be negated completely. Whether they persist in altered form or re-emerge in similar form after the pandemic comes under control, the implementation of teacher evaluation and teacher collaboration policies will present coherence challenges that administrators and teachers will have to manage, most likely through related structures and interacting responses. Thus, while the recognition of this study's contribution to our knowledge of teacher quality policy implementation may be delayed, it may not, I trust, be indefinitely denied.

APPENDICES

APPENDIX A

Focus Teacher/PLC Member Survey

In accordance with your informed consent form, your name, the name of your school, the name of your district, and your responses to the following questions will be kept strictly confidential. You may skip any question, or part of a question, that you do not wish to answer; however, we hope that you will candidly and fully answer as many questions as possible. Thank you.

Name: _____ **Date: _____**

For each item, please indicate your choice or answer by putting the appropriate number (1-4 or 1-5) from the stated scale in the appropriate blank or by filling in the appropriate blank or space, as indicated in the item.

1. **Pleases indicate the level of effectiveness formal teacher evaluation provides with respect to each of the following**:
   Please fill in each blank with the appropriate number:
   1 = No or very low effectiveness, 2 = low effectiveness, 3 = moderate effectiveness, 4 = high effectiveness, 5 = very high effectiveness

_____ Formative assessment for teachers' professional learning and improvement
_____ Summative assessment of teachers' effectiveness and quality
_____ Summative assessments for employment and staffing decisions (for example, retention, tenure, lay-off, dismissal)
_____ Motivation for teachers to improve
_____ Compliance with district and state mandates
_____ Other *(please specify):* _____

2. **Please indicate the level of influence each of the following has on how you participate in your formal evaluation:**
   Please fill in each blank with the appropriate number:
   1 = No or very low influence, 2 = low influence, 3 = moderate influence, 4 = high influence, 5 = very high influence

_____ The need to comply with evaluation requirements and expectations
_____ Your evaluator relationship and interactions
_____ Your informal interactions with other teachers and professional friends
_____ The nature and value of the feedback you receive
_____ Your personal experiences and beliefs
_____ Your grade/course level PLC tasks, experiences, and interactions
_____ Other *(please specify):* _____.

3. **Pleases indicate the level of effectiveness grade/course level PLCs provide with respect to each of the following**:

Please fill in each blank with the appropriate number:

1 = No or very low effectiveness, 2 = low effectiveness, 3 = moderate effectiveness, 4 = high effectiveness, 5 = very high effectiveness

_____ Authentically collaborative teacher professional learning and improvement
_____ Development of effective curricula, classroom practices and assessments
_____ Formation of collective norms and practices for student learning
_____ Motivation for teachers to improve
_____ Compliance with school and district mandates
_____ Other *(please specify):* _____

4. **Please indicate the level of influence each of the following has on your participation in your grade/course level PLC:**

Please fill in each blank with the appropriate number:

5 = No or very low influence, 2 = low influence, 3 = moderate influence, 4 = high influence, 5 = very high influence

_____ The need to comply with PLC requirements and expectations
_____ Your PLC relationships and interactions
_____ The nature and value of your PLC work
_____ Your informal interactions with other teachers and professional friends
_____ Your personal experiences and beliefs
_____ Your formal evaluation requirements, experiences and interactions
_____ Other *(please specify):* _____

5. **Please indicate the level of influence each of the following has on your how your grade/course level PLC functions as a group:**

Please fill in each blank with the appropriate number:

1 = No or very low influence, 2 = low influence, 3 = moderate influence, 4 = high influence, 5 = very high influence

_____ The goals, tasks and expectations set for your grade/course level PLC's work
_____ The administrative oversight and supervision of your grade/course level PLC's work
_____ The individual personalities and characteristics of your grade/course level PLC's members
_____ The collective interactions and learning that occur during your grade/course level PLC's work
_____ The need of all your grade/course level PLC members to complete their own formal evaluation
_____ Other *(please specify):* _____

6. Please indicate the extent to which you agree or disagree with each of the following statements:
Please fill in each blank with the appropriate number:
1 = Strongly Disagree, 2 = Disagree, 3 = Agree, 4 = Strongly Agree

_____ Formal evaluation does not significantly affect my collaboration or other interactions with my grade/course level PLC's members.
_____ Formal evaluation does not significantly affect the conduct or results of my grade/course level PLC's work.
_____ Formal evaluation does not significantly affect the meaningfulness or usefulness of my grade/course level PLC's work.
_____ Formal evaluation does not put significantly more pressure for accountability on me than my grade/course level PLC's work.
_____ Formal evaluation does not motivate me to improve significantly more than my grade/course level PLC's work.

7. Please indicate the extent to which you agree or disagree with each of the following statements:
Please fill in each blank with the appropriate number:
1 = Strongly Disagree, 2 = Disagree, 3 = Agree, 4 = Strongly Agree

_____ I care more about my formal evaluation than I do about my grade/course level PLC's work.
_____ I apply more attention and effort to my formal evaluation than to my grade/course level PLC's work.
_____ The results from the time, attention and effort I devote to my formal evaluation are worth more to me than the results from the time, attention and effort I devote to my grade/course level PLC's work.
_____ The requirement that all teachers undergo annual formal evaluations seems to encourage competition among teachers, including those in my grade/course level PLC.

8. Including the current school year, how many years have you been a teacher?
Please fill in the blanks. _____ years **In this district? _____** years **In this school? _____** years

9. *In the space provided*, **please indicate your current teaching assignment.**
_____
_____
_____
_____

10. Have you taught *all* the classes in your current assignment before the current school year?
Please fill in the appropriate blank(s).

_____ Yes

_____ No—*please specify the percentage or proportion as a fraction of your assignment that involves class(es) you have* not *taught before the current school year:* _____
_____.

11. *In the blank provided,* **please indicate the main academic department you belong to in this school during the current school year. _____**

12. *In the blank provided*, **please indicate the main grade level PLC you belong to in your main department in this school during the current school year. _____**

13. What is your highest post-secondary degree? *Please put an "X" in the appropriate blank.*

_____ Bachelors
_____ Masters
_____ Education Specialist
_____ Doctorate
_____Other highest education degree (*Please specify*: _____)

14. With what gender do you identify? *Please put an "X" in the appropriate blank.*

_____ Female
_____ Male
_____ Non-binary/other identification

15. What is your race/ethnicity? *Please put an "X" in the appropriate blank.*

_____ African American/Black
_____ Caucasian/White
_____ American Indian/Native American
_____ Hispanic/Latinx
_____ Asian or Pacific Islander
_____ Other or multiple race/ethnicity

16. Please indicate your age. *Please fill in the blank.* _____ years

This concludes the questionnaire.
Thank you for your time, effort and thoughtful participation.

APPENDIX B
Focus Teacher/PLC Member First Interview Protocol

[Note: The information that answers much of Question 1 below is covered by the focus teacher questionnaire.]

1. Please give me some background about your teaching in the district and high school:
    a. How long have you been a teacher? Taught in the district? At the high school?
    b. What is your current assignment? Have you had the same or similar assignment in the past? What is your main department and grade level PLC?
    c. What is it like teaching in your school? What are the main challenges? How would you describe the school's climate and culture?

2. Please describe your most recent (last year's) annual evaluation from start to finish. How did your evaluation go, and how did it work out for you?
   Possible probes/prompts:
    a. What do you like and dislike about it? How might it be improved?
    b. How would you describe your relationship and interactions with your evaluator?
    c. What do you think of the time, attention and effort you devoted to your evaluation? How would you describe its value or usefulness?
    d. Did the evaluation process and result go as you expected?
    e. What, if any, difference in your evaluation process have you experienced or do you expect this year?

3. What influences how you approach and participate in the evaluation process? Could you give me an example of each influence?
   Possible probes/prompts:
    a. The goals, components and requirements of the process?
    b. Your prior experiences and beliefs?
    c. Your informal interactions with other teachers and professional friends?
    d. Administrative leadership and support?
    e. Your grade level PLC relationships and work?
    f. Your participation in district/school events—e.g., staff meetings, other PLC times, other professional development activities?

4. Has your participation in the evaluation process affected how you work or interact with other teachers and administrators? If so, could you give me an example?
   Possible probes/prompts: An example from:
    a. how you get or give help, make professional decisions, or enact your own practice?
    b. your informal collegial interactions?
    c. your interactions with your grade level PLC?
    d. your interactions in formal work settings, such as professional development activities, staff meetings, department meetings, etc.?

5. Please describe your most recent (last year's) grade level PLC? How did your grade level PLC's work go, and how did it work out for you?
   Possible probes/prompts:

a. What did you like and dislike about it? How might it be improved?
b. How would you describe your relationships and interactions with your grade level PLC?
c. What do you think of the time, attention and effort you devoted to your grade level PLC? How would you describe its value or usefulness of your grade level PLC work?
d. Did your grade level PLC's work and result go as you expected?
e. What, if any, difference in your grade level PLC work and interactions have you experienced or do you anticipate this year?

6. What influences how you approach and participate in your grade level PLC? Could you give me an example of each influence?
   Possible probes/prompts:
   a. The required goals, tasks and processes?
   b. Your experiences and beliefs?
   c. Your informal interactions with other teachers and professional friends?
   d. Administrative leadership and support?
   e. Your participation in the evaluation process?
   f. Your participation in district/school events—e.g., staff meetings, other PLC times, professional development activities?

7. Could you tell me about a time when your work with one—evaluation or grade level PLC—connected or overlapped with your work in the other? For example, a time when:
   Possible probes/prompts:
   a. You applied what you learned from one to the work of other?
   b. You used what you produced in one in the work of the other?
   c. You brought up something from one in a discussion of the other?
   d. You received help in one and shared it in the other?

8. Do you think that how you participate in the evaluation process and how you participate with your grade level PLC affect or relate to each other in some way? If so, could you give me an example?
   Possible probes/prompts:
   a. How your approach or mindset toward one affects your approach or mindset toward the other?
   b. How your cooperation or competition in one affects your cooperation or competition in the other?
   c. How your goals for one affect your goals for the other?
   d. How you allocate, direct or redirect, your time, attention and effort from one to the other?

9. Thinking about how evaluations and grade level PLCs now work and have previously worked in your school, does the way one is or has been done or help or hinder your current participation in the other? Could you give me an example of how one helps or hinders the other?
   Possible probes/prompts:

a. Does the way one is or has been done affect individual or group professional learning in the other?
b. Does the way one is or has been done affect competition or cooperation in the other?
c. Does the way one is or has been done affect individual or group compliance in the other?
d. Does the way one is done or has been done affect how time, attention and energy are devoted to or spent on the other?

10. How would describe the overall relationship between your responsibilities, goals, work, and expectations concerning your evaluation and those of your grade level PLC?
Possible probes/prompts:
    a. Completely separate and independent?
    b. Consistent and complementary?
    c. Opposed and conflicting?
    d. Other description/characterization?

11. Concluding questions:
    a. Beyond what we've discussed, what else is important for me to understand about your participation in the evaluation process or your grade level PLC?
    b. Would you like to add or clarify anything about the topics we've discussed?

APPENDIX C
Focus Teacher/PLC Member Second Interview Protocol

1. First Semester Generally
    a. Overall how did the first semester go for you?
    b. Any particular personal/professional goals or expectations for the second semester?

2. Evaluation
    a. Did you have any personal/professional goals or expectations for the evaluation process this past semester? What did you want or hope to get out of it?
    b. What is the current status of your evaluation—that is, what steps have been completed?
    c. How has it gone for you so far?
        i. Did anything happen that has affected how you carried out your job roles/duties (unusual, unexpected, challenging, surprising)?
        ii. Did anything happen that has affected your prior attitude or beliefs about evaluation, including its value or usefulness?
        iii. Did anything happen that has affected how you participate in your grade-level PLC, interact with your PLC colleagues or make use of your PLC work?
    d. What steps are left to complete in your evaluation in the second semester? Do you have any particular goals or expectations around its completion?

3. Grade-Level PLC during Departmental Content Time
    a. Did you have any personal/professional goals or expectations for your grade-level PLC work this past semester? What did you want or hope to get out of it?
    b. What is the current status of your grade-level PLC work—that is, what steps have been completed?
    c. How has it gone for you so far?
        i. Did anything happen that has affected how you carried out your job roles/duties (unusual, unexpected, challenging, surprising)?
        ii. Did anything happen that has affected your prior attitude or beliefs about grade-level PLC work, including its value or usefulness?
        iii. Did anything happen that has affected how you participate in your evaluation, interact with your evaluator or make use of some aspect of your evaluation?
    d. How would you describe how your grade-level PLC worked as a group during the first semester? Generally and in terms of:
        i. Achieving or satisfying PLC goals/expectations?
        ii. Collaboration and professional learning?
        iii. Collegiality and collegial support?
    e. Do you think your grade-level PLC tasks, interactions or results were affected by the evaluation process that was going for individual teachers?
    f. What is left to complete in your grade-level PLC work in the second semester? Do you have any particular goals or expectations around its completion?

4. Evaluation/PLC Connections?
    a. On reflection, do you see any connection or relationship between your evaluation and grade-level PLC work in how you think about, participate in, or carry out your obligations in each?
    b. Was any connection or relationship between the two required or encouraged by administration?
    c. Did you or PLC develop any connection or relationship?
    d. Since the goals and potential value of evaluation and PLCs lie in teacher learning and development, if weak or nonexistent relationship or connection, why? How might each be modified or improved to develop positive and useful connections?

5A. [Dist/HS #1 Only] Collaboration Time PLC
    a. Which such PLC did you belong first semester? Second semester?
    b. How would you compare the work, interactions and value of the two kinds of PLC in which you participate? Which more worthwhile to you, participating in which would you rather spend time?
    c. Do you see any connection or relationship between your evaluation and collaboration-time PLC work in how you think about, participate in, or carry out your obligations in each?

5B. [Dist/HS #2 Only] Mid-Year Review Meeting
    a. What can you tell me about the "Mid-Year Review," which I understand is a new thing this year? For example, who/what is involved, what is its purpose, how is it being conducted, and with what result?
        [Ask about any documentation, including a list of potential questions, any other materials, for preparation, review or other use during any meeting.]
    b. What is the status of your own mid-year review preparation, meeting or results? That is, what's been done and what's left to do with respect to it? How did it go, or how do you expect it to go? Why?
    c. What, if any connection, was given or do you perceive between the mid-year review and your evaluation? The review and your PLC work?
    d. Has or might the mid-year review affect how you approach or participate in your evaluation during the rest of the year? Same question for your PLC work for the rest of the year?

6. Concluding Question
    a. Beyond what we've discussed, what else is important for me to understand about your participation in the evaluation process or your grade level PLC? Would you like to add or clarify anything about either interview or any observation?
    b. Do you have any questions or suggestions for me?

APPENDIX D
Building Administrator (Principal, Assistant Principal) Interview Protocol

1. Confirm name of administrator and position/title:
    a. How long have you served in your current position?
    b. How would you describe your current position, including its main roles and responsibilities (especially with respect to teacher evaluations and PLCs)?

2. How does teacher evaluation work in your school?
   Possible probes/prompts:
    a. What are your goals for teacher evaluation?
    b. What decisions do you make about teacher evaluation, and what roles do building administrators play in teacher evaluation?
    c. How do you provide leadership and support for your teachers' evaluations?
    d. How do building administrators use the results of teachers' evaluations?

3. What are the challenges and effects of teacher evaluation in your school?
   Possible probes/prompts:
    a. How would you describe a successful teacher evaluation?
    b. What are the main concerns your teachers have about evaluation, and how does building administration address those concerns?
    c. How do you think evaluation affects your teachers? Your relationships with your teachers? Your school's professional culture?
    d. How might teacher evaluation be improved?

4. How do content time PLCs, including grade level PLCs within departments, work in your school?
   Possible probes/prompts:
    a. What are your goals for these PLCs?
    b. What decisions do you make about these PLCs, and what roles do building administrators play in these PLCs?
    c. How do you provide leadership and support for these PLCs?
    d. How do building administrators use the results of the work of these PLCs?

5. What are the challenges of content time PLCs in your school?
   Possible probes/prompts:
    a. How would you describe a content time PLC that works effectively?
    b. What are the main concerns your teachers have about these PLCs, and how do building administrators address those concerns?
    c. How do you think these PLCs affect your teachers? Your relationships with your teachers? Your school's professional culture?
    d. How might the work of these PLCs be improved?

6. What connection, if any, exists between the conduct of teacher evaluations and the work of these PLCs in your school?

Possible probes/prompts:

a. Is there an explicitly expected or mandated connection?
b. Is there a custom or practice that makes a connection?

7. Do you think teachers in your school see or experience a connection between teacher evaluation and content time PLCs? If so, how would you describe it with a couple of examples?
    a. Can you describe an occasion when something about a teacher's content time PLC work came up during their evaluation?
    b. Can you describe an occasion when a teacher's participation in their evaluation seemed affected by their content time PLC work?
    c. Can you describe an occasion when something about their evaluation came up during a teacher's content time PLC work?
    d. Can you describe an occasion when a teacher's participation in their content time PLC seemed affected by their evaluation?

8. How would you compare teacher evaluations and content time PLCs in your school:
    a. Do you see or experience any connection between your roles in each?
    b. How would you compare the benefits of each to your teachers? Why?
    c. How would you compare the significance of the impact of each on teachers' work or performance? The school's culture and performance? Why?

9. Concluding questions:
    a. Beyond what we've discussed, what else is important for me to understand about your school's teacher evaluations or content time PLCs?
    b. Do you wish to add or clarify anything about the topics we've discussed?

APPENDIX E
District Central Office Administrator Interview Protocol

1. Confirm name of administrator and position/title:
    a. How long have you served in your current position?
    b. How would you describe your current position, including its main roles and responsibilities (especially with respect to teacher evaluations and PLCs)?

2. Please describe when and how your district adopted and implemented:
    a. the current teacher evaluation system.
    b. the current system of PLCs, particularly content time departmental PLCs with grade level PLCs.

3. How does teacher evaluation work in your district's high schools?
   Possible probes/prompts:
    a. What are the main goals for teacher evaluation?
    b. What are the main components of the teacher evaluation system, and who is responsible for completing each component?
    c. How does your district use the results of high school teacher evaluation?
    d. What, if anything, is different about teacher evaluation this school year?

4. What does your district do to provide leadership and support for high school teacher evaluation?
   Possible probes/prompts:
    a. For building administrators?
    b. For teachers?

5. What challenges does high school teacher evaluation present to your district?
   Possible probes/prompts:
    a. How do you think the teacher evaluation affects your district's high school? Its administrators, teachers, professional culture?
    b. How might teacher evaluation be improved?

6. How do content time PLCs work in your district's high school?
   Possible probes/prompts:
    a. What are the goals of high school PLCs? Assigned tasks?
    b. How are high school PLCs organized and run?

7. What does your district do to provide leadership and support for high school content time PLCs?
   Possible probes/prompts:
    a. How are content time high school PLCs overseen or supervised?
    b. What, if anything, is different about high school content time PLCs this school year?

8. What challenges do high school content time PLCs present?

Possible probes/prompts:

a. How do you think these PLCs affects your district's high school? Its administrators, teachers, professional culture?
b. How might these PLCs be improved?

9. What connection, if any, exists in the district's high schools between teacher evaluation and content time PLCs?
   Possible probes/prompts:
   a. Is there an explicitly expected or mandated connection? If so, could you give an example of how the connection works?
   b. Is there a custom or practice that makes a connection? If so, could you give an example of how the connection works?

10. How would you compare the effects of teacher evaluations and content time PLCs in your district's high schools?
    a. How would you compare the benefits to high school teachers'? Central or building administrators? Why?
    b. What would you change about one to align or improve the functioning and outcomes of the other? Why?

11. Concluding questions:
    a. Beyond what we've discussed, what else is important for me to understand about your district's systems of teacher performance evaluation or content time PLCs at the high school level?
    b. Would you like to add or clarify anything we've discussed?

APPENDIX F
Code Descriptions

- **Institutionalized Evaluation Structure**: with respect to the state-mandated system of annual teacher performance evaluation, its formal components, relationships, rules and resources; and its norms, processes and practices as implemented by the case high schools in their particular districts.

    - Evaluation as district implemented: how the district of a case high school establishes, organizes and operates its institutionalized evaluation structure; excludes structural linkages (defined below) between evaluation and PLC structures.

    - Evaluation as school implemented: how a case high school establishes, organizes and operates its institutionalized evaluation structure; excludes structural linkages (defined below) between evaluation and PLC structures.

- **Evaluation Work Determination**: with respect to the content or "what" of teacher participation in their own evaluation, who primarily decides the objectives, tasks, activities and products to be engaged in by the teacher.

    - Teacher-driven evaluation work: evaluation objectives, tasks, activities, and products that are primarily determined by individual teachers as compared to administrators.

    - Administration-driven evaluation work: evaluation objectives, tasks, activities, and products that are primarily determined by administrators as compared to individual teachers.

- **Teacher Evaluation Participation Purposes**: how teachers and administrators explain why teachers participate in their evaluation; includes descriptions of teacher goals, rationales, intentions, and motivations for participation.

    - Evaluation Compliance: teachers participate to obey or show adherence to evaluation demands; includes expressions such as "because we have to," "it's required," and "there's no choice," and "it's a hoop to jump through"; also includes a desire to avoid adverse consequences such as a low summative rating or the desire for a high final summative rating.

    - Professional learning to inform/improve practice: teachers participate to obtain formative feedback or other resources from their administrator-evaluator concerning their evaluation, which the teacher may then use to modify or enhance their own teaching.

    - Supporting school or professional norms or culture: teachers participate to uphold or follow collective expectations for appropriate teacher attitudes, beliefs and behaviors in their case high school; includes norms and practices around teacher job status, duties or responsibilities; being accepted or recognized as "professional," a "good teacher," and doing "a good job"; also includes such norms and practices as teacher collaboration, equality, autonomy, and professional expertise and judgment.

- **Influences on Evaluation Participation**: teacher beliefs and experiences that significantly inform or affect their attitude toward or enactment of their evaluation; excludes teacher beliefs and experiences related to structural linkages between evaluation and PLCs (defined below).

    - PLC-related influences: general teacher beliefs about or experiences with PLCs that inform or affect their attitude toward or enactment of their evaluation; includes discussion in PLC related to evaluation, but unrelated to relevant PLC work or goals, and excludes instances where it is the tasked, assigned or expected PLC work or its products that affect evaluation participation.

    - Non-PLC-related influences: any influence that is not related to general teacher beliefs about or experiences with PLCs; includes influences stemming from beliefs about or experiences with, for example, personality traits, and teacher preparation or induction; and, apart from evaluation, professional learning activities, informal peer interactions, and administrative leadership or support.

- **Value of Evaluation Participation**: descriptions of the significance, utility, meaningfulness, or other value of evaluation to teachers, including any of its activities, products or effects; includes expressions about the value of evaluation overall or of any evaluation component or process.

    - Positive value evaluation participation: favorable descriptions of the significance, utility, meaningfulness, or other value of evaluation; includes positively valued evaluation components and processes, as well as favorable changes in evaluation.

    - Neutral or mixed value evaluation participation: neutral (neither explicitly positive or negative) or combined positive and negative descriptions of the significance, utility, meaningfulness, or other value of evaluation; includes neutrally or combined valued evaluation components and processes, as well as neutral or combined expressions about changes in evaluation; also includes expressions such as "it doesn't concern me," or "I don't worry about it."

    - Negative value evaluation participation: unfavorable descriptions of the significance, the utility, meaningfulness, or other value of evaluation; includes negatively valued evaluation components and processes, unfavorable changes in evaluation.

- **Institutionalized PLC Structure**: with respect to the model of systematic teacher collaboration in formalized professional learning communities (PLCs) that a district with a case high school and the case high school adopts, the model's formal components, relationships, rules and resources; and its norms, processes and practices as implemented by the case high schools in their particular districts.

    - PLCs as district implemented: how the district of a case high school establishes, organizes and operates its institutionalized PLC structure; excludes structural linkages (defined below) between evaluation and PLC structures.

    - PLCs as school implemented: how a case high school establishes, organizes and operates its institutionalized PLC structure; excludes structural linkages (defined below) between evaluation and PLC structures.

- **PLC Work Determination**: with respect to the content or "what" of teacher participation in their PLC, who decides upon PLC objectives, tasks, activities and products to be engaged in by teachers.

    - Teacher-driven PLC work: PLC objectives, tasks, activities, and products that are primarily determined by a PLC's teacher-members as compared to administrators.

    - Administration-driven PLC work: PLC objectives, tasks, activities, and products that are primarily determined by administrators as compared to a PLC's teacher-members.

- **Teacher PLC Participation Purposes**: why teachers participate in their PLCs; includes teacher descriptions of their goals, rationales, intentions, and motivations for participation.

    - PLC Compliance: teachers participate to obey or show adherence to PLC demands; includes expressions such as "because we have to," "it's required," and "there's no choice," and "it's a hoop to jump through," as well as to avoid any adverse consequences of non-compliance.

    - Professional learning to inform/improve practice: teachers participate for professional learning from and in collaboration with their PLC peers, which the teacher may then apply to modify or enhance their teaching and student learning.

    - Supporting school or professional norms or culture: teachers participate to uphold or follow collective expectations for appropriate teacher attitudes, beliefs and behaviors in their case high school; includes norms and practices around teacher job status, duties or responsibilities; being accepted or recognized as "professional," a "good teacher," and doing "a good job"; also includes such norms and practices as teacher collaboration, equality, autonomy, and professional expertise and judgment.

- **Influences on PLC Participation**: those teacher beliefs and experiences that significantly inform or affect their attitudes toward or enactment of their PLC work; excludes teacher beliefs and experiences related to structural linkages between evaluation and PLCs (defined below).

    - Evaluation-related influences: general teacher beliefs about or experiences with evaluation inform or affect their attitude toward or enactment of PLC; excludes instances where it is required or expected evaluation components or their products that affect PLC participation.

    - Non-evaluation-related influences: any influence that is not related to general teacher beliefs about or experiences with evaluation; includes influences stemming from beliefs about or experiences with, for example, personality traits, and teacher preparation or induction; and, apart from PLCs, professional learning activities, informal peer interactions, and administrative leadership or support.

- **Value of PLC Participation**: teacher descriptive assessments of the significance, utility, meaningfulness, or other value of PLCs to their teacher-members, including any of its activities, products or effects; includes expressions about the value of PLCs overall or of any PLC

component or process.

- Positive value PLC participation: favorable descriptions of the significance, utility, meaningfulness, or other value of PLCs; includes positively valued PLC components and processes, as well as favorable changes in PLCs.

- Neutral or mixed value evaluation participation: neutral (neither explicitly positive or negative) or combined positive and negative descriptions of the significance, utility, meaningfulness, or other value of PLCs; includes neutrally or combined valued PLC components and processes, as well as neutral or combined expressions about changes in PLCs; also includes expressions such as "it doesn't concern me," or "I don't worry about it."

- Negative value evaluation participation: unfavorable descriptions of the significance, the utility, meaningfulness, or other value of PLCs; includes negatively valued PLC components and processes, unfavorable changes in PLCs.

- **Evaluation/PLC Relationship and Inter-Relationship**: how evaluation and PLCs compare or relate to each other, plus those interactions or interconnections that provide opportunities common to or shared by both structures through which teachers may take their response to the demands of one and adapt their response to the demand of the other.

  - Teacher/administrator perceptions/experience of the evaluation/PLC relationship: how teachers and administrators describe the relationship between evaluation and PLCs, including any relationship or connection between their respective components and processes; includes descriptions of how they have experienced any relationship, the absence of any relationship, and descriptions of any comparison or contrast.

  - Structural linkages: the nature, form, and properties of those opportunities for teacher adaptive response strategies arising from the structural interactions or interconnections between evaluation and PLCs; linkages encourage, incentivize or facilitate teacher adaptive response strategies, whether developed or initiated by teachers or administrators.

  - Teacher responses to structural linkages: how teachers engage with a structural linkage through the use of an adaptive response strategy whereby their response to the demands of one structure also serves their response to the demands of the other.

REFERENCES

REFERENCES

Achinstein, B. (2002). Conflict Amid Community: The Micropolitics of Teacher Collaboration. *Teachers College Record, 104*(3), 421-455.

Adams, J. E., Jr., and Kirst, M. W. (1999). New Demands and Concepts for Educational Accountability: Striving for Results in an Era of Excellence. In J. Murphy and K. S. Louis (eds.), *Handbook of Research on Educational Administration* (pp. 463-489). Washington, DC: AERA.

Akiba, M. and Liang, G. (2016). Effects of teacher professional learning activities on student achievement growth. *The Journal of Educational Researcher, 109*(1), 99-110.

Andrews-Larson, C., Wilson, J., and Larbi-Cherif, A. (2017). Instructional Improvement and Teachers' Collaborative Conversations: The Role of Focus and Facilitation. Teachers College Record, 119(2), 1-36.

Banerjee, N., Stearns, E., Moller, S., Mickelson, R. A. (2016). Teacher Job Satisfaction and Student Achievement: The Roles of Teacher Professional Community and Teacher Collaboration in Schools. *American Journal of Education, 123*(February), 203-241.

Baxter, P. and Jack, S. (2008). Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers, *The Qualitative Report, 13*(4), 544-559.

Bazeley, P. (2013). *Qualitative Data Analysis: Practical Strategies.* Thousand Oaks, CA: SAGE Publications, Inc.

Bernard, H. R., Wutich, A., and Ryan, G. W. (2017). *Analyzing Qualitative Data: Systematic Approaches,* 2nd ed. Thousand Oaks, CA: SAGE Publications, Inc.

Blankenship, S. S. and Ruona, W. E. A. (2007). Professional Learning Communities and Communities of Practice: A Comparison of Models, Literature Review. *Online Submission,* ERIC Resources in Education (ED 504776). Retrieved from https://eric.ed.gov/?id=ED504776.

Bolhuis, E., Schildkamp, K., and Voogt, J. (2016). Data-based decision making in teams: enablers and barriers. *Educational Research and Evaluation, 22*(3-4), 213-233.

Brodersen, M., Joyce, J., Bopp, L., Stewart, J., and Cherasaro, T. (2017). *District implementation of educator evaluation systems: Survey results.* Report prepared for the Michigan Department of Education. Centennial, CO: Marzano Research, retrieved from https://www.michigan.gov/documents/mde/Marzano_Research-district_implementation_of_educator_evaluation_sytems-survey_results_605554_7.pdf.

Brouwer, P., Brekelmans, M., Nieuwenhuis, L., and Simons, R. (2012). Fostering teacher community development: A review of design principles and a case study of an innovative interdisciplinary team. *Learning Environment Research, 15,* 319-344.

Bryk, A. S., Gomez, L. M., Grunow, A., and LeMahieu, P. G. (2016). *Learning to Improve: How America's Schools Can Get Better at Getting Better.* Cambridge, MA: Harvard Education Press.

Bryk, A. S. and Schneider, B. (2002). *Trust in Schools: A Core Resource for Improvement.* New York, NY: Russell SAGE Foundation.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., and Easton, J. Q. (2010). *Organizing Schools for Improvement: Lessons from Chicago.* Chicago, IL: The University of Chicago Press.

Buchanan, R. (2015). Teacher identity and agency in an era of accountability. *Teachers and Teaching: theory and practice, 21*(6), 700-719.

Burch, P. (2007). Educational Policy and Practice from the Perspective of Institutional Theory: Crafting a Wider Lens. *Educational Researcher, 36*(2), 84-95.

Cameron-Standerford, A., Anderson, D., Bergh, B., Selling, J., Bohjanen, S., and Barnes, A. (2017). *Michigan Teachers' and Administrators' Perceptions of the Teacher Evaluation Process.* Northern Michigan University-Michigan Department of Education report, retrieved from https://www.michigan.gov/documents/mde/NMUMichiganTeacherEvaluationProcess_121117FINAL-adaready_611410_7.pdf.

Buttram, J. L. and Farley-Ripple, E. N. (2016). The Role of Principals in Professional Learning Communities. *Leadership and Policy in Schools, 15*(2), 192-220.

Charner-Laird, M., Ng, M., Johnson, Moore Johnson, S., Kraft, M. A., Papay, J. P., and Reinhorn, S. K. (2016). Gauging Goodness of Fit: Teachers' Assessments of their Instructional Teams in High-Poverty Schools. Working Paper, Project of the Next Generation of Teacher, Harvard Graduate School of Education. Retrieved from https://projectngt.gse.harvard.edu/files/gse-ojectngt/files/gauging_goodness_of_fit_0622916.pdf.

Chingara, R. and Heystek, J. (2019). Leadership as agency in the context of structure. *International Journal of Educational Management, 33*(7), 1596-1609.

Corbin, J. and Strauss, A. (2008*). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory,* 3rd ed. Thousand Oaks, CA: SAGE Publications, Inc.

Coburn, C. E. (2001). Collective Sensemaking about Reading: How Teachers Mediate Reading Policy in Their Professional Communities. *Educational Evaluation and Policy Analysis, 23*(2), 145–170.

Coburn, C. E. (2016). What's Policy Got to Do with It? How the Structure-Agency Debate Can Illuminate Policy Implementation. *American Journal of Education, 122*(May), 465-475.

Conley, S. and Enomoto, E. K. (2005). Routines in school organizations: Creating stability and change. *Journal of Educational Administration, 43*(1), 9-21.

Creswell, J. W. (2013). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches,* 3rd ed. Thousand Oaks, CA: SAGE Publications, Inc.

Danielson, C. (2007). *Enhancing Professional Practice: A Framework for Teaching,* 2nd ed. Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L. (1989). Accountability for Professional Practice. *Teachers College Record, 91*(1), 59-80.

Darling-Hammond, L. (1990). Teacher Professionalism: Why and How? In A. Lieberman (ed.), *Schools as Collaborative Cultures.* School Development and the Management and Change Series: 3. Bristol, PA: The Falmer Press.

Darling-Hammond, L. (2013). *Getting Teacher Evaluation Right: What really matters for effectiveness and improvement.* New York, NY: Teachers College Press.

Darling-Hammond, L. and McLaughlin, M. W. (1999). Investing in Teaching as a Learning Profession: Policy Problems and Prospects. In L. Darling-Hammond and G. Sykes (eds.*), Teaching as the Learning Profession: Handbook of Policy and Practice* (pp. 315-340). San Francisco: Jossey-Bass, A Wiley Imprint.

Darling-Hammond, L., Wise, A. E., and Pease, S. R. (1983). Teacher Evaluation in the Organizational Context: A Review of the Literature. *Review of Educational Research, 53*(3), 285-328.

Datnow, A. (2011). Collaboration and contrived collegiality: Revisiting Hargreaves in the age of accountability. *Journal of Educational Change, 12,* 147-158. DOI 10.1007/s10833-011-9154-1.

Debray, E., Parson, G., and Avila, S. (2003). Internal Alignment and External Pressure: High School Responses in Four State Contexts. In Carnoy, M., Elmore, R., and Siskin, L. S. (eds.), *The New Accountability: High Schools and High-Stakes Testing* (pp. 55-86). New York, NY: RoutledgeFalmer.

Dee, T. S. and Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management, 34*(2), 267-297.

Desimone, L. M. (2013). Teacher and Administrator Responses to Standards-Based Reform. *Teachers College Record, 115*(1), 1-53.

DiPaola, M. F. and Tschannen-Moran, M. (2005). The impact of schools' adaptive strategies on student achievement. *Journal of Educational Administration, 43*(1), 60-71.

Donaldson, M. L. (2009). *So Long Lake Woebegon? Using Teacher Evaluation to Raise Teacher Quality.* Washington, D.C.: Center for American Progress, retrieved from https://www.americanprogress.org/issues/education-k-12/reports/2009/06/25/6243/so-long-lake-wobegon/

Donaldson, M. L. (2012). *Teachers' Perspectives on Evaluation Reform.* Washington, D.C.: Center for American Progress, retrieved from https://www.americanprogress.org/issues/education-k-12/reports/2012/12/13/47689/teachers-perspectives-on-evaluation-reform/

Donaldson, M. L. and Mavrogordato, M. (2018). Principals and teacher evaluation: The cognitive, relational, and organizational dimensions of working with low-performing teachers. Journal of Educational Administration, 56(6), 586-601.

Donaldson, M. L. and Papay, J. P. (2014). Teacher Evaluation for Accountability and Development. In H. F. Ladd and M. E. Goertz (eds.), *Handbook of Research in Education Finance and Policy* (pp. 174-190). ProQuest Ebook Central: Routledge.

Donaldson, M. L. and Woulfin, S. (2018). From Tinkering to Going "Rogue": How Principals Use Agency When Enacting New Teacher Evaluation Systems. *Educational Evaluation and Policy Analysis, 40*(4), 531-556.

Donohoo J. (2018). Collective teacher efficacy research: Productive patterns of behavior and other positive consequences. *Journal of Educational Change*, *19*, 323-345.

Doss, C. J. and Akinniranye, G. (2019). School Reform Efforts: Here Today, Gone Tomorrow? Educator Perspectives on the Rapid Life Cycle of School Reforms. Rand Data Note: Insights from the American Educator Panels, retrieved from https://www.rand.org/pubs/research_reports/RR2575z3-1.html.

DuFour, Richard. (2004). What Is a "Professional Learning Community"? *Educational Leadership, 61*(8, May), 6-11.

DuFour, Richard., DuFour, Rebecca., Eaker, R., and Karhanek, G. (2004). *Whatever It Takes: How Professional Learning Communities Respond When Kids Don't Learn.* Bloomington, IN: Solution Tree.

DuFour, Richard. and Eaker, R. (1998). *Professional Learning Communities at Work: Best Practices for Enhancing Student Achievement.* Bloomington, IN: Solution Tree (formerly National Educational Service).

DuFour, Richard. and Fullan, M. (2013). *Cultures Built to Last: Systemic PLCs at Work.* Bloomington, IN: Solution Tree Press.

Dulude, E., Spillane, J. P., and Xavier Dumay, X. (2017). High Stakes Policy and Mandated

Curriculum: A Rhetorical Argumentation Analysis to Explore the Social Processes That Shape School Leaders' and Teachers' Strategic Responses. *Educational Policy, 31*(3), 364-403.

Elmore, R. (2005). Accountable Leadership. *The Educational Forum, 69*(Winter), 134-142.

Elmore, R., Ableman, C. H., Even, J., Kenyon, S., and Marshall, J. (2004). When Accountability Knocks, Will Anyone Answer? In R. Elmore (ed.), *School Reform from the Inside Out: Policy, Practice, and Performance* (pp. 133-199). Cambridge, MA: Harvard Education Press.

Emma White Research (2019). Launch Michigan: 2019 Educator Survey, retrieved from https://www.launchmichigan.org/wp-content/uploads/2019/03/Launch-MI-Full-Report-2019-FINAL.pdf.

Erichsen, K. and Reynolds, J. (2020). Public school accountability, workplace culture, and teacher morale. *Social Science Research, 85,* 1-15. Retrieved from https://doi.org/10.1016/j.ssreach.2019.102347.

Erickson, F. (1986). Qualitative research in education. In M. C. Wittrock (Ed.), *Handbook of research on teaching,* 3rd ed. (pp. 119-161). Washington, DC: American Educational Research Association.

Feger, S. and Arruda, E (2008). *Professional Learning Communities: Key Themes from the Literature.* Report of The Education Alliance at Brown University (Spring 2008). Retrieved from https://www.brown.edu/academics/education-alliance/publications/professional-learning-communities-key-themes-literature.

Feldman, M. S. and Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly, 48*(1), 94-118.

Finnigan, K. S., and Gross, B. (2007). Do Accountability Policy Sanctions Influence Teacher Motivation? Lessons from Chicago's Low-Performing Schools. *American Educational Research Journal, 44*(3), 594-629.

Firestone, W. A. (1993). Alternative arguments for generalizing from data as applied to qualitative research. *Educational Researcher*, *22*(4), 16-23.

Firestone, W. A. and Shipps, D. (2005). How Do Leaders Interpret Conflicting Accountabilities to Improve Student Learning? In W. A. Firestone and C. Riehl (eds.), *A New Agenda for Research in Educational Leadership* (pp. 81-100). New York: Teachers College Press.

Forsythe, P. B., Adams, C. M., and Hoy, W. K. (2011). *Collective Trust: Why Schools Can't Improve Without It.* New York, NY: Teachers College Press.

Fullan, M. and Hargreaves, A. (2015). Professional Capital as Accountability. *Education Policy Analysis Archives, 23*(15), 1-17.

Fuhrman, S. H. (2003a). Preface. In S. H. Fuhrman (ed.), *Designing Coherent Education Policy: Improving the System* (pp. xi-xx). San Francisco, CA: Jossey-Bass Publishers.

Fuhrman, S. H. (2003b). The Politics of Coherence. In S. H. Fuhrman (ed.), *Designing Coherent Education Policy: Improving the System* (pp. 1-34). San Francisco, CA: Jossey-Bass Publishers.
Glesne, C. (2016). *Becoming Qualitative Researchers: An Introduction,* 5th ed. New York, NY: Pearson Education, Inc.

Goe, L., Bell, C., and Little, O. (2008). *Approaches to Teacher Effectiveness: A Research Synthesis.* Report of the National Comprehensive Center for Teacher Quality (June 2008). Retrieved from https://gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf.

Goetz, J.P. and LeCompte, M. D. (1984). *Ethnography and qualitative design in educational research.* New York, NY: Academic Press.

Gordon, K. (2008). Tightening the Ship or Slowly Sinking? Reshaping Teachers' Work Conditions. In B. Fuller, M. K. Henne, and E. Hannum (eds.), *Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability* (pp. 31-66). Research in Sociology of Education, Volume 16, B. Fuller and E. Hannum (eds.). Wagon Lane, Bingley, UK: JAI Press, Imprint of Emerald Group Publishing, Ltd.

Guest, G., MacQueen, K. M., and Narney, E. E. (2014). *Applied Thematic Analysis: Themes and Codes*. SAGE Research Methods. Thousand Oaks, CA: SAGE Publications, Inc.

Hallinger, P., Heck, R. H., and Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Education Assessment and Evaluation Accountability*, DOI 10.1007/s11092-013-9179-5, 5-23.

Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age.* New York, NY: Teachers College Press.

Hargreaves, A. (2007). Sustainable professional learning communities. In L. Stoll and K. S. Louis (eds.), *Professional Learning Communities: Divergence, Depth, Dilemmas* (Chap 13, pp. 181-195). Berkshire, England: Open University Press, McGraw-Hill Education.

Hargreaves, A. and Dawe, R. (1990). Paths of Professional Development: Contrived Collegiality, Collaborative Culture, and the Case of Peer Coaching. *Teaching and Teacher Education, 6*(3), 227-241.

Harris, D. N. and Herrington, C. D. (2015). Editors' Introduction: The Use of Teacher Value-Added Measures in Schools: New Evidence, Unanswered Questions, and Future Prospects. *Educational Researcher, 44*(2), 71-76.

Hess, F. M. (1999). *Spinning Wheels: The Politics of Urban School Reform*. Washington, DC: Brookings Institution Press.

Hewitt, K. K. (2015). Educator Evaluation Policy that Incorporates EVAAS Value-Added Measures: Undermined Intentions and Exacerbated Inequities. *Education Policy Analysis Archives, 23*(76), 1-35.

Honig, M. I. and Hatch, T. C. (2004). Crafting Coherence: How Schools Strategically Manage Multiple, External Demands. *Educational Researcher, 33*(8), 16-30.

Hord, S. (1997). *Professional Learning Communities: Communities of Continuous Inquiry and Improvement.* Report of the Southwest Educational Development Laboratory (now American Institutes for Research [AIR]). Retrieved from https://sedl.org/pubs/change34/.

Horn, I., Garner, B., Chen, I-C. and Frank, K. A. (2020). Seeing Colleagues as Learning Resources: The Influence of Mathematics Teacher Meeting Quality on Advice-Seeking Social Networks. AERA Open, 6(2), 1-19. Retrieved from https://journals.sagepub.com/doi/full/10.1177/2332858420914898.

Horn, I. S., Garner, B., Kane, B. D., and Brasel, J. (2017). A Taxonomy of Instructional Learning Opportunities in Teachers' Workgroup Conversations. *Journal of Teacher Education, 68*(1), 41-54.

Horn, I. S., Kane, B. D., and Garner, B. (2018). Teacher Collaborative Time: Helping Teachers Make Sense of Ambitious Teaching in the Context of Their Schools. In P. Cobb, K. Jackson, E. Henrick, T. M. Smith, and the MIST Team (eds.), *Systems for Instructional Improvement: Creating Coherence from the Classroom to the District Office* (Chap 6, pp. 93-112). Cambridge, MA: Harvard Education Press.

Horn, I. S. and Little, J. W. (2010). Attending to Problems of Practice: Routines and Resources for Professional Learning in Teachers' Workplace Interactions. *American Educational Research Journal, 47*(1), 181-217.

Ingersoll, R.M. (2003). *Who Controls Teachers' Work? Power and Accountability in America's Schools.* Cambridge, MA: Harvard University Press.

Ingersoll, R. M. and Collins, G. J. (2017). Accountability and Control in American Schools. *Journal of Curriculum Studies, 49*(1), 75-95.

Jiang, J. Y., Sporte, S. E., and Lupppescu, S. (2015). Teacher Perspectives on Evaluation Reform: Chicago's REACH Students. *Educational Researcher, 44*(2), 105-116.

Johnson, B. (2003). Teacher Collaboration: good for some, not so good for others. Educational Studies, 29(4), 337-351

Johnson, S. M. (2015). Will VAMS Reinforce the Walls of the Egg-Crate School*? Educational Researcher, 44*(2), 117-126.

Johnson, S. M., Kraft, M. A., and Papay, J. P. (2012). How Context Matters in High-Need Schools: The Effects of Teachers' Working Conditions on Their Professional Satisfaction and Their Students' Achievement. *Teachers College Record, 114*(10), 1-35.

Johnson, S. M., Reinhold, S., and Simon, N. (2018). Ending Isolation: The Payoff of Teacher Teams in Successful High-Poverty Urban Schools. *Teachers College Record, 120*(5), 1-46.

Jones, N. D., Bettini, E., and Brownell, M. T. (2016). Competing Strands of Educational Reform Policy: Can Collaborative School Reform and Teacher Evaluation Reform Be Reconciled? Albert Shanker Institute research report, retrieved from http://www.shankerinstitute.org/resource/competingstrands.

Kelley, C. J. and Finnigan, K. (2003). The Effects of Organizational Context on Teacher Expectancy. *Educational Administration Quarterly, 39*(5), 603-634.

Kimball, S. M. (2002). Analysis of Feedback, Enabling Conditions and Fairness Perceptions of Teachers in Three School Districts with New Standards-Based Evaluation Systems. *Journal of Personnel Evaluation in Education, 16*(4), 241-268.

Knapp, M. S., Bamburg, J. D., Ferguson, M. C., and Hill, P. T. (1998). Converging Reforms and the Working Lives of Frontline Professionals in Schools. *Educational Policy, 12*(4), 397-418.

Knapp, M. S. and Feldman, S. B. (2012). Managing the intersection of internal and external accountability: Challenge for urban school leadership in the United States. *Journal of Educational Administration, 50*(5), 666-694.

Koppich, J. E. and Kerchner, C. T. (1999). Organizing the Other Half of Teaching. In L. Darling-Hammond and G. Sykes (Eds.*), Teaching as the Learning Profession: Handbook of Policy and Practice* (pp. 315-340). San Francisco: Jossey-Bass, A Wiley Imprint.

Kraft, M. A. and Gilmour, A. F. (2016). Can Principals Promote Teacher Development as Evaluators? A Case Study of Principals' Views and Experiences. *Education Administration Quarterly, 52*(5), 711-753.

Kraft, M. A. and Gilmour, A. F. (2017). Revisiting *The Widget Effect:* Teacher Evaluation Reforms and the Distribution of Teacher Effectiveness. *Educational Researcher, 46*(5), 234-249.

Kraft, M. A. and Papay, J. P. (2014). Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Evaluation and Policy Analysis, 36*(4), 476-500.

Kruse, S. D., Louis, K. S., and Bryk, A. S. (1995). An Emerging Framework for Analyzing School-Based Professional Community. In K. S. Louis, S. D. Kruse, and Associates (eds.), *Professionalism and Community: Perspectives on Reforming Urban Schools* (Chap 2, pp. 23-42). Thousand Oaks, CA: Corwin Press, Inc., a SAGE Publications Company.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37-62.

Lane, J. L. (2020). Maintaining the Frame: Using Frame Analysis to Explain Teacher Evaluation Policy Implementation. *American Educational Research Journal, 57*(1), 5-42.

Lavigne, A. L. (2014). Exploring the Intended and Unintended Consequences of High-Stakes Teacher Evaluation on Schools, Teachers, and Students. *Teachers College Record, 116*(1), 1-29.

Lee, V. E. and Smith, J. B. (1996). Collective Responsibility for Learning and Its Effects on Gains in Achievement for Early Secondary School Students. *American Journal of Education, 104* (February), 103-136.

Lenhof, S. (2012). Strengthening Michigan's Teacher Force: How a new teacher evaluation system will better equip Michigan educators to improve student achievement. The Education Trust-Midwest, retrieved from https://www.bridgemi.com/wp-content/uploads/2012/09/73387_ETI_WidgetEffect.pdf.

Levine, T. H. (2010). Tools for the Study and Design of Collaborative Teacher Learning: The Affordances of Different Conceptions of Teacher Community and Activity Theory. *Teacher Education Quarterly 37*(1), 109-130.

Lipsky, M. (2010 [1980]). *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services.* 30th Anniversary Ed. New York: Russell Sage Foundation.

Little, J. W. (1990). The Persistance of Privacy: Autonomy and Initiative in Teachers' Professional Relations. *Teachers College Record, 91*(4), 509-536.

Little, J. W. (2002a). Locating learning in teachers' communities of practice opening up problems of analysis in records of everyday work. *Teaching and Teacher Education, 18*, 917-946.

Little, J. W. (2002b). Professional community and the problem of high school reform. *International Journal of Educational Research, 37*, 693-714.

Loeb, S., Miller, L. C., and Wyckoff, J. (2015). Performance screens for school improvement: The case of teacher tenure reform in New York City. *Educational Researcher, 44*(4), 199-212.

Lomos, C., Hofman, R. H., and Bosker, R. J. (2011). Professional communities and student achievement – a meta-analysis. *School Effectiveness and School Improvement, 22*(2), 121-148.

Lortie, D. C. (2002 [1975]). *Schoolteacher*. 2d ed. Chicago, IL: The University of Chicago Press.

Louis, K. S., Dretzke, B., and Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness and School Improvement, 21*(3), 315-336.

Louis, K. S., Kruse, S. D., and Associates (eds.) (1995*). Professionalism and Community: Perspectives on Reforming Urban Schools.* Thousand Oaks, CA: Corwin Press, Inc., a SAGE Publications Company.

Louis, K. S., Kruse, S. D., and Bryk, A. S. (1995). Professionalism and Community: What Is It and Why Is It Important in Urban Schools? In K. S. Louis, S. D. Kruse, and Associates (eds.), *Professionalism and Community: Perspectives on Reforming Urban Schools* (Chap 1, pp. 3-22). Thousand Oaks, CA: Corwin Press, Inc., a SAGE Publications Company.

Louis, K. S. and Lee, M. (2016). Teacher capacity for organizational learning: the effects of school culture and context. *School Effectiveness and School Improvement, 27*(4), 534-556.

Louis, K. S. and Marks, H. M. (1998). Does Professional Community Affect the Classroom? Teachers' Work and Student Experiences in Restructuring Schools. *American Journal of Education, 106* (August), 532-565.

Louis, K. S., Marks, H. M., and Kruse, S. (1996). Teachers' Professional Community in Restructuring Schools. *American Educational Research Journal, 33*(4), 757-798.

Louis, K. S. and Robinson, V. M. (2012). External mandates and instructional leadership: school leaders as mediating agents. *Journal of Educational Administration, 50*(5), 629-665.

Marsh, J. A., Bush-Mecenas, S., Strunk, K. O., Lincove, J. A., and Huguet, A. (2017). Evaluating Teachers in the Big Easy: How Organizational Context Shapes Policy Reponses in New Orleans. *Educational Evaluation and Policy Analysis, 39*(4), 539-570.

Marzano Research (2017). *Teacher and Administrator Evaluation Systems: Case Study Report.* Centennial, CO: Marzano Research. Report prepared for the Michigan Department of Education, retrieved from https://www.michigan.gov/documents/mde/MarzanoResearchFinalCaseStudyReport508-Update_610904_7.pdf.

Maslow, V. J. and Kelley, C. J. (2012). Does Evaluation Advance Teaching Practice? The Effects of Performance Evaluation on Teaching Quality and System Change in Large Diverse High Schools. *Journal of School Leadership, 22*, 600-630.

Maxwell, J. A. (2013). *Qualitative Research Design: An Interactive Approach*, 3rd ed. No. 41 of Applied Social Research Methods Series. Thousand Oaks, CA: SAGE Publications, Inc.

McGuinn, P. (2012). Stimulating reform: Race to the top, competitive grants, and the Obama education agenda. *Educational Policy, 16,* 136-159.

McLaughlin, M. W. and Talbert, J. E. (2001). *Professional Communities and the Work of High School Teaching.* Chicago, IL: The University of Chicago Press.

Mehta, J. (2013). *The Allure of Order: High Hopes, Dashed Expectations, and the Troubled Quest to Remake American Schooling.* New York: Oxford University Press.

Mich. Comp. Laws §1248 (2018).

Mich. Comp. Laws §1249 (2018).

Mich. Comp. Laws §1249(1) (2018).

Mich. Comp. Laws §1249(1)(d) (2018).

Mich. Comp. Laws §1249(2)(a) (2018).

Mich. Comp. Laws §1249(2)(f) (2018).

Mich. Comp. Laws §1249(2)(j) (2018).

Mich. Comp. Laws §1249(2)(k) (2018).

Michigan Council for Educator Effectiveness (2013). Building an Improvement-Focus System of Educator Evaluation in Michigan: Final Recommendations, retrieved from http://www.mcede.org/system/resources/BAhbBlsHOgZmSSI7MjAxMy8wNy8yMy8yMS80NC81NS83NzYvbWlkbmlnaHRyZXBvcnRfanVseTI0XzIwMTMucGRmBjoGRVQ/midnightreport_july24_2013.pdf.

Michigan Department of Education (nd.a). Educator Evaluations and Effectiveness in Michigan: An analysis of 2013-14 educators evaluation systems survey and educator effectiveness data, retrieved from https://www.michigan.gov/documents/mde/2013-14_Educator_Evaluations_and_Effectiveness_485909_7.pdf.

Michigan Department of Education (nd.b). *Michigan Educator Evaluations At-a-Glance.* Office of Educator Excellence, retrieved from https://www.michigan.gov/documents/mde/Educator_Evaluations_At-A-Glance_522133_7.pdf.

Michigan Department of Education (2019). MI School Data, retrieved from https://www.mischooldata.org/.

Miesner, H. R., Blair, E. E., Packard, C. C., Velaquez, M., Macgregor, L., and Grodsky, E. (2019). Collaborating in Context: Relational Trust and Collaborative Structures in Eight Wisconsin Elementary Schools. WCER Working Paper No. 2019-7 (September 2019). Wisconsin Center for Education Research, University of Wisconsin-Madison. Retrieved from https://wcer.wisc.edu/docs/working-papers/Working_Paper_No_2019_7.pdf.

Miles, M. B., Huberman, A. M., and Saldana, J. (2014). *Qualitative data analysis: A Methods Sourcebook,* 3rd Ed. Thousand Oaks, CA: SAGE Publications, Inc.

Møller, J. (2008). School leadership in an age of accountability: Tensions between managerial and professional accountability. *Journal of Educational Change, 10,* 37-46.

Moran, R. M. R. (2017). The Impact of a High Stakes Teacher Evaluation System: Educator Perspectives on Accountability. *Educational Studies, 53*(2), 178-193.

Murphy, J., Hallinger, P., and Heck, R. H. (2013). Leading via Teacher Evaluation: The Case of the Missing Clothes? *Educational Researcher, 42*(6), 349-354.

Murphy, J. and Torre, D. (2014). *Creating Productive Cultures in Schools for Students, Teachers, and Parents.* Thousand Oaks, CA: Corwin, A SAGE Company.

National Council on Teacher Quality (2017). *2017 State Teacher Policy Yearbook.* Retrieved from https://www.nctq.org/publications/2017-State-Teacher-Policy-Yearbook.

Newmann, F. M., King, M. B., and Rigdon, M. (1997). Accountability and School Performance: Implications from Restructuring Schools. *Harvard Educational Review, 67*(1), 41-69.

Newmann, F. M., Smith, B., Allensworth, E., and Bryk, A. S. (2001). Instructional Program Coherence: What It Is and Why It Should Guide School Improvement Policy. *Education Evaluation and Policy Analysis, 23*(4), 297-321.

O'Day, J. A. (2002). Complexity, Accountability, and School Improvement. *Harvard Educational Review, 72*(3), 292-329.

O'Day, J. A. (2008). NCLB and the Complexity of School Improvement. In A. R. Sadovnik, J. A. O'Day, G. W. Bohrnstedt, and K. M. Borman (eds.), *No Child Left Behind and the Reduction of the Achievement Gap* (pp. 25-52). New York, NY: Routledge, Taylor & Francis Group.

Papay, J. P. (2012). Refocusing the Debate: Assessing the Purposes and Tools of Teacher Evaluation. *Harvard Educational Review, 82*(1), 123-141.

Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods: Integrating Theory and Practice,* 4th ed. Thousand Oaks, CA: SAGE Publications, Inc.

Pogodzinski, B., Umpstead, R., and Witt, J. (2015). Teacher evaluation reform implementation and labor relations. *Journal of Education Policy, 30*(4), 540-561.

Peterson, P. L. and Comeaux. M. A. (1990). Evaluating the Systems: Teachers' Perspectives on Teacher Evaluation. *Educational Evaluation and Policy Analysis, 12*(1), 3-24.

Popp, J. S., and Goldman, S. R. (2016). Knowledge building in teacher professional learning communities: Focus of meeting matters. *Teacher and Teacher Education, 59*, 347-359.

Printy, S. M. (2008). Leadership for Teacher Learning: A Community of Practice Perspective. *Educational Administration Quarterly, 44*(2), 187-226.

Public Employment Relations Act, Mich. Comp. Laws §§423.201, et seq.

Quintero, E. (ed.) (2017). *Teaching in Context: The Social Side of Education Reform.* Cambridge, MA: Harvard Education Press.

Ray Tayor and Associates (2017). *Educator Evaluation Research and Evaluation Activities Final Report.* Report to the Michigan Department of Education, retrieved from https://www.michigan.gov/documents/mde/FINAL_MDE_Educator_Evaluation_Final_Report_605652_7.pdf.

Reid, D. B. (2020). Teachers' perceptions of how principals use new teacher evaluation systems. *Teachers and Teaching, 26*(1), 129-144.

Reinhorn, S., Johnson, S. M., and Simon, N. S. (2017). Investing in Development: Six High-Performing, High-Poverty Schools Implement Massachusetts' Teacher Evaluation Policy. *Educational Evaluation and Policy Analysis, 39*(3), 383-406.

Reinhorn, S. and Johnson, S. M. (2015). Can Teacher Evaluation Provide Both Accountability and Development? Learning from Six Schools' Implementation of Evaluation Policy. Project on the Next Generation of Teachers, Harvard Graduate School of Education, retrieved from https://projectngt.gse.harvard.edu/files/gse-projectngt/files/can_evaluation_provide_accountability_and_development_1016.pdf.

Revised School Code of 1976, as amended by PA 2009, No. 205, Imd. Eff. Jan. 4, 2010, codified as Mich. Comp. Laws §380.1249, further amended by PA 2010, No. 336, Imd. Eff. Dec. 21, 2010; PA 2011, No. 102, Imd. Eff. July 19, 2011; PA 2015, No. 173, Imd. Eff. Nov. 5, 2015; PA 2018, No. 235, Eff. Sept. 29, 2018.

Richmond, G. and Manokore, V. (2010). Identifying Elements Critical for Functional and Sustainable Professional Learning Communities. *Science Teacher Education,* DOI 10.1002/sce.20430. Retrieved from www.wileyonlinelibrary.com.

Rigby, J. G. (2014). Principals' sensemaking and enactment of teacher evaluation. Journal of *Educational Administration, 53*(3), 385-390.

Rigby, J. G., Larbi-Cherif, A., Rosenquist, B. A., Sharpe, C. J., Cobb, P., and Smith, T. (2017). Administrator Observation Feedback: Does It Lead Toward Improvement in Inquiry-Oriented Math Instruction? *Educational Administration Quarterly*, DOI: 10.1177/0013161X16687006, 1-42.

Rigby, J. G., Woulfin, S. L., and Marz, V. (2016). Understanding How Structure and Agency Influence Education Policy Implementation and Organizational Change. *American Journal of Education, 122*(May), 295-301.

Robinson, S. (2012). Constructing teacher agency in response to the constraints of education policy: adoption and adaptation. *The Curriculum Journal, 23*(2), 231-245.

Robinson, V., Bendikson, L., McNaughton, S., Wilson, A., and Zhu, T. (2017). Joining the Dots: The Challenge of Creating Coherent School Improvement. *Teachers College Record, 119*(9), 1-43.

Ronfeldt, M., McQueen, K., and Grissom, J. A. (2015). Teacher Collaboration in Instructional Teams and Student Achievement. *American Educational Research Journal, 52*(3), 475-514.

Rowan, B. (1990). Commitment and Control: Alternative Strategies for the Organizational Design of Schools. *Review of Research in Education, 16,* 353-359.

Rowan, B., Schilling, S. G., Spain, A., Bhandari, P., Berger, D., and Graves, J. (2013). Promoting High Quality Teacher Evaluations in Michigan: Lessons from a Pilot of Educator Effectiveness Tools. Institute for Social Research, University of Michigan, retrieved from http://mcede.org/system/resources/BAhbBlsHOgZmSSJXMjAxNC8wMi8wNC8yMC80NS8wNy80NzcvUHJvbW90aW5nX0hpZ2hfUXVhbGl0eV9UZWFjaGVyX0V2YWx1YXRpb25zX2luX01pY2hpZ2FuLnBkZgY6BkVU/Promoting-High-Quality-Teacher%20Evaluations-in-Michigan.pdf.

Russell, J. L. and Bray, L. E. (2013). Crafting Coherence from Complex Policy Messages: Educators' Perceptions of Special Education and Standards-Based Accountability Policies. *Education Policy Analysis Archives, 21*(12), 1-18.

Scott, W. R. (2008). *Institutions and Organizations: Ideas and Interests.* 3rd ed. Thousand Oaks, CA: Sage.

Serdiouk, M., Bopp, L., and Cherasaro, T. (2017). *Teacher Perceptions of Evaluator Feedback: Final Report.* Report prepared for the Michigan Department of Education. Centennial, CO: Marzano Research, retrieved from https://www.michigan.gov/documents/mde/Marzano_Research_Teacher_Perceptions_O_fEvaluator_Feedback_Final_Report1_611115_7.pdf.

Sewell, W.H. (1992). A Theory of Structure: Duality, Agency, and Transformation. *American Journal of Sociology, 98*(1), 1-29.

Shirrell, M. (2016). New principals, accountability, and commitment in low-performing schools. *Journal of Educational Administration, 54*(5), 558-574.

Silver, C. and Lewins, A. (2014). *Using Software in Qualitative Research: a step-by-step guide,* 2nd ed. Thousand Oaks, CA: SAGE Publications, Inc.

Spillane, J. P., Reiser, B. J., and Reimer, T. (2002). Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. *Review of Educational Research, 72*(3), 387-431.

Sporte, S. E., Stevens, W. D., Healey, K., Jiang, J., and Hart, H. (2013). *Teacher Evaluation in Practice: Implementing Chicago's REACH Students.* Research Report (September 2013),

University of Chicago Consortium on Chicago School Research. Retrieved from https://consortium.uchicago.edu/publications/teacher-evaluation-practice-implementing-chicagos-reach-students.

Srinivasan, L. and Archer, J. (2018). *From Fragmentation to Coherence: How more integrative ways of working could accelerate improvement and progress.* Report of the Carnegie Corporation of New York (November 2018). Retrieved from https://www.carnegie.org/topics/topic-articles/improving-schools-systems/accelerating-improvement-education-equity-solutions-address-fragmentation/.

Stake, R. E. (2005). Qualitative case studies. In N. Denzin & Y. Lincoln (eds.), *Handbook of Qualitative Research,* 2nd ed (pp. 443-466). Thousand Oaks, CA: SAGE Publishing.
Stake, R. E. (2006). *Multiple Case Study Analysis.* New York, New York: Guilford Press.

Stone-Johnson, C. (2015). Intensification and isolation: Alienated teaching and collaborative professional relationships in the accountability context. *Journal of Educational Change, 17,* 29-39.

Stosich, E. L. (2018). Principals and teachers "craft coherence" among accountability policies. *Journal of Educational Administration, 56*(2), 203-219.

Stoll, L., Bolam, R., McMahon, A., Wallace, M., and Thomas, S. (2006). Professional Learning Communities: A Review of the Literature. *Journal of Educational Change, 7*, 221-258.

Stone-Johnson, C. (2016). Intensification and Isolation: Alienated teaching and collaborative professional relationship in the accountability context. *Journal of Educational Change, 17*, 29-49.

Supovitz, J. A. (2002). Developing Communities of Instructional Practice. *Teachers College Record, 104*(8), 1591-1626.

Supovitz, J., Sirinides, P., and May, H. (2010). How Principals and Peer Influence Teaching and Learning. *Educational Administration Quarterly, 46*(1), 31-56.

Sykes, G. (1999). The "New Professionalism" in Education: An Appraisal. In J. Murphy and K. S. Louis (eds.), *Handbook of Research on Educational Administration* (pp. 227-249). Washington, DC: AERA.

Szczesiul, S. A. and Huizenga, J. L. (2015). Bridging Structure and Agency: Exploring the Role of Teacher Leadership in Teacher Collaboration. *Journal of School Leadership, 25,* 368-405.

Talbert, J. E. (2010). Professional Learning Communities at the Crossroads: How Systems Hinder or Engender Change. In A. Hargreaves, et al. (eds.), *Second International Handbook of Educational Change* (pp. 555-571). Springer Science and Business Media B.V.

Talbert, J. E. and McLaughlin, M. W. (2008). Teaching Quality as a Problem of School Change. In A. R. Sadovnik, J. A. O'Day, G. W. Bohrnstedt, and K. M. Borman (eds.), *No Child Left Behind and the Reduction of the Achievement Gap* (pp. 177-200). New York, NY: Routledge, Taylor & Francis Group.

Talbert, J. E. and McLaughlin, M. W. (1994). Teacher Professionalism in Local School Contexts. *American Journal of Education, 102*(2), 123-153.

Taylor, E. S. and Tyler, J. H. (2012). The Effect of Evaluation on Teacher Performance. *American Economic Review, 102*(7), 3628-3651.

Teacher Tenure Act, Mich. Comp. Laws §§38.71, et seq.

Thiel, C., Schweigher, S., and Bellmann, J. (2017). Rethinking Side Effects of Accountability in Education: Insights from a Multiple Methods Study in Four German School Systems. *Education Policy Analysis Archives, 25*(93), 1-27.

Tichnor-Wagner, A., Harrison, C., and Cohen-Vogel, L. (2016). Cultures of Learning in High Schools. *Educational Administrative Quarterly, 52*(4), 602-642.

Tuytens, M. and Devos, G. (2010). The influence of school leadership on teachers' perception of teacher evaluation policy. *Educational Studies, 36*(5), 521-536.

Valli, L. and Buese, D. (2007). The Changing Roles of Teachers in an Era of High-Stakes Accountability. *American Educational Research Journal, 44*(3), 519-558.

Valli, L., Croninger, R. G., and Walters, K. (2007). Who (Else) is the Teacher? Cautionary Notes on Teacher Accountability Systems. *American Journal of Education, 113* (August), 635-662.

Vescio, V., Ross, D., and Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education, 24*, 80-91.

Weatherley, R. and Lipsky, M. (1977). Street-Level Bureaucrats and Institutional Innovation: Implementing Special Education Reform. *Harvard Educational Review, 47*(2), 171-197.

Weathers, J. M. (2014). Teacher Community in Urban Elementary Schools: The Role of Leadership and Bureaucratic Accountability. Education Policy Analysis Archives, *19*(3), 1-33.

Weisberg, D., Sexton, S., Mulhern, J., and Keeling, D. (2009). *The Widget Effect: Our National Failure to Acknowledge and Act on Differences in Teacher Effectiveness*, 2nd ed. Report of The New Teacher Project, retrieved from https://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf.

White, K. W. and Rosenbaum, J. E. (2008). Inside the Black Box of Accountability: How High Stakes Accountability Alters School Culture and the Classification and Treatment of Students

and Teachers. In A. R. Sadovnik, J. A. O'Day, G. W. Bohrnstedt, and K. M. Borman (eds.), *No Child Left Behind and the Reduction of the Achievement Gap* (pp. 177-200). New York, NY: Routledge, Taylor & Francis Group.

Whitehurst, G. J., Chingos, M. M., and Lindquist, K. M. (2014). *Evaluating Teachers with Classroom Observations: Lessons Learned in Four Districts.* Report of Brown Center on Education Policy at Brookings. Retrieved from https://www.brookings.edu/research/evaluating-teachers-with-classroom-observations-lessons-learned-in-four-districts/.

Wills, J. S. and Sandholtz, J. H. (2009). Constrained Professionalism: Dilemmas of Teaching in the Face of Test-Based Accountability. *Teachers College Record, 111*(4), 1065-1114.

Wood, D. (2007). Teachers' Learning Communities: Catalyst for Change or a New Infrastructure for the Status Quo? *Teachers College Record, 109*(3), 699-739.

Woodland, R. H. and Mazur, R. (2015). Beyond Hammers Versus Hugs: Leveraging Educator Evaluation and Professional Learning Communities into Job-Embedded Professional Development. *NASSP Bulletin*, *99*(1), 5-25.

Woulfin, S. L. (2016). Duet or Duel? A Portrait of Two Logics of Reading Instruction in an Urban School District. *American Journal of Education, 122,* 337-362.

Woulfin, S. L., Donaldson, M. L., and Gonzales, R. (2016). District Leaders' Framing of Educator Evaluation Policy. *Educational Administration Quarterly, 52*(1), 110-135.

Yin, R. K. (2014). *Case study research: Design and Methods,* 5th ed. Thousand Oaks, CA: SAGE Publications, Inc.

Young, V. M. (2006). Teachers' Use of Data: Loose Coupling, Agenda Setting, and Team Norms. *American Journal of Education, 112,* 521-548.