

Examining How School Settings Support Teachers' Improvement of their Classroom Instruction

By  
Charlotte Jean Dunlap

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in  
Learning, Teaching, and Diversity

August, 2016

Nashville, Tennessee

Paul A. Cobb, Ph.D.

Melissa Gresalfi, Ph.D.

Deborah Rowe, Ph.D.

Thomas M. Smith, Ph.D.

## ACKNOWLEDGEMENTS

The research reported here was supported in part by the National Science Foundation, grant no. ESI-0554535 and no. DRL-0830029. The opinions expressed are those of the author and do not necessarily reflect the views of the National Science Foundation.

This dissertation would not have been possible without the support of many individuals. First, I am very grateful to Dr. Paul Cobb, who has been a key support in the dissertation process as well as my own development as a researcher and writer. His commitment to improving teaching and learning in public schools is inspiring, and an encouraging reminder that the research-practice gap is only as wide as we choose to make it.

I would also like to thank my committee members, Dr. Melissa Gresalfi, Dr. Deborah Rowe, and Dr. Tom Smith for their advice, encouragement, and pushback over the last two years. They have helped me become a more well-rounded researcher and a more effective writer, and have taught me how to speak and listen in larger circles than my immediate one.

I consider myself exceptionally fortunate to have worked on the MIST project with such a diverse, caring, supportive, and intelligent group of PIs, graduate students, research assistants, and postdocs over the last five years. Through my experiences on MIST, I have learned the value of teamwork, of collaborating with others who see the world differently from me, and of being a critical friend and mentor to others. To Dr. Annie Wilhelm, Dr. Lynsey Gibbons, and Dr. Erin Henrick especially, your advice and mentorship were critical to my development.

To Brooks Rosenquist, who has been an incredible friend, collaborator, and source of accountability and support through the dissertation process, I owe many thanks. The contributions of his that I name in the methods section of this dissertation represent only a small

portion of his role in supporting my growth as a researcher. To Amy Farris and Nicholas Kochmanski—you have been wonderful sources of support, wisdom, and laughter, and I look forward to our continued friendships and collaborations.

To my family, I am especially indebted. My parents' love and encouragement has been constant, and their love of learning has always been an inspiration. My two amazing sisters, Amanda and Megan, have been essential sources of love and support. And finally, I am grateful to my remarkable husband Brent Sharpe. He is a constant model of decency, openness, and strength; without his unwavering support, I would not be the person I am today.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	vi
LIST OF FIGURES .....	vii
 Chapter	
I: Introduction .....	1
Background & Motivation.....	4
Research questions .....	4
Brief Study Overview.....	8
Definition of Key Terms.....	9
II: Review of the relevant literature .....	12
Prior research relating teachers’ knowledge, vision, beliefs, and practices .....	12
Prior research on the influence of professional development on teachers’ practices .....	18
Prior research on the influence of instructional coaching on teachers’ practices.....	21
Prior research on the influence of teacher collaboration on teachers’ practices.....	25
Prior research on the influence of principal leadership on teachers’ practices.....	29
Summary .....	32
III: Methodology.....	34
Research Design.....	34
Addressing RQ 1.....	36
Addressing RQ2.....	37
Research setting: Districts B and D.....	38
District Vision. ....	39
Key District Strategies. ....	40
Changes beginning in year 5.....	42
Participants .....	45
Data and Measures .....	47
Assessing teachers’ instructional expertise and identification.....	48
Analyzing aspects of teachers’ school settings. ....	54
Case Selection .....	55
Potential measurement error and hidden variables in case selection and analysis. ....	61
Analysis: Framework and Methods.....	62
Conceptual Framework.....	62
Analysis Methods.....	67
Limitations.....	73
IV: Findings .....	75
Introduction .....	75
Part 1: School-Based Supports and Expectations for Improvers’ Development.....	77
One-On-One Work with Instructional Coaches .....	78
Pull-out professional development. ....	85
District PD initiatives.....	86
Findings.....	87

Teacher Collaborative time (TCT).....	92
School leaders' Instructional Expectations and Vision.....	96
Part 2: Instructional Expertise, Identification with Reform Practices, and Instructional Improvement .....	99
Diane and Cathy (improvers) and Roger (stable) at Cypress Middle School.....	100
Diane and Cathy: accessing support to improve.....	104
Mario (improver) and Lois (stable) at Two Lagoons.....	113
Mario: supports for improvement. ....	117
Lois: supports for her learning. ....	124
Summary .....	126
Anna (improver) and Roxanne (stable) at Park Falls.....	127
Anna: supports for improvement .....	128
Synthesis .....	136
V: Conclusions & Implications.....	139
Critical aspects of teachers' current practice.....	139
Critical Supports for Teacher Learning.....	141
Implications for Future Research in Teacher Learning.....	145
Implications for Research and Practice in Professional Development .....	146
Developing an ambitious vision.....	147
Anticipating challenges and developing productive views of diverse students..	147
References .....	152
Appendix	
A. Codebook for A-Priori coding .....	162
B. Memo Template for Case Summary Form.....	163
C. Buy-in/Identification Memo Template.....	164
D. Improver vs. Stable Teacher Memo Template .....	165
E. Triangulation memo template 1 .....	166
F. Triangulation memo template 2 .....	167
G. Instructional Expertise of Case Teachers .....	168
H. Instructional Expertise of Case Teachers' School Leader(s).....	169

## LIST OF TABLES

Table	Page
1. Summary of Research Design .....	38
2. Descriptive statistics for IQA data in districts B and D, years 1-7.....	57
3. Summary of Cases Selected for Analysis.....	59
4. Principles of High-Quality Supports for Teacher Learning .....	65
5. Descriptive Information for all Cases .....	77
6. Summary of teachers' participation in math-specific PD.....	87
7. Case Summaries: school, case type, years, and grade levels. ....	99

## LIST OF FIGURES

	Page
1. Roger's year 2 spring survey responses regarding coaching .....	111
2. Roger's year 2 spring survey responses regarding CMP2 .....	112

## **Chapter I**

### **Introduction**

The recent widespread adoption of the Common Core State Standards for Mathematics (CCSS-M; National Governors Association for Best Practices & Council of Chief State School Officers, 2010), which emphasize students' conceptual understanding in mathematics along with their development of procedural fluency, presents a significant challenge to teachers. Supporting students to develop domain-appropriate practices in mathematics entails new instructional practices that differ greatly from what most teachers experienced when they were students. For example, supporting students to “construct viable arguments and critique the reasoning of others” (CCSS-M, MP3) implicates the work of teachers in selecting and implementing classroom tasks that afford students the opportunity to reason, justify, and make sense of others' ideas in everyday instruction. This poses a substantial challenge to teachers who have historically been held accountable for their students' mastery of skills and procedures, with little attention to their conceptual understanding of mathematics. Not only do more rigorous standards implicate the need for teachers to further develop their own understanding of the content, but they also presume a vision of high-quality instruction (Munter, 2014) that many teachers have not developed or personally experienced as students (Cohen, 2011; Hiebert et al., 2005).

If common supports for teacher learning (e.g., pull-out professional development, the ongoing support of an instructional coach, school-based collaborative time) are to effectively help teachers develop practices aligned with more rigorous standards, these supports must be organized around the aspects of teachers' knowledge and practice that are integral to ambitious

teaching. A number of research studies have been conducted which document teachers' learning in the context of such PD – e.g., recurring sessions organized around particular aspects of teachers' knowledge (Elliott et al., 2009), vision (Sherin & van Es, 2009), and practice (Nelson, Lesseig, Slavit, Kennedy, & Seidel, 2015).

However, even in cases where teachers engage in well-designed opportunities to learn, these forms of support may yet be insufficient for supporting teachers to learn more ambitious practices *and enact them* in their everyday instruction (Feiman-Nemser, 2001; Olson and Barrett, 2004). Researchers in teacher education have documented a persistent “two worlds” problem: the value systems, goals for student learning, vision of instruction, and repertoire of tools and practices shared among those in the teacher education setting are often vastly different from those present in the school communities that novice teachers enter into (Campbell, 2012; Gainsburg, 2012; Horn, Nolen, Ward, and Campbell, 2008; Thompson, Windschitl, and Braaten, 2013). As a result of these competing systems or worlds, teachers must negotiate their membership or affiliation between the two in ways that often lead to inconsistencies between practice across settings. As Horn et al., (2008) point out, “the repertoire of practices that [teachers] develop, and their justification for using or not using them, can largely be traced to these two worlds [of teacher development settings and their field experiences]” (p. 6). A similar disconnect between teachers' discourse and practice in in-service professional development (PD) settings and their classroom practice has been the focus of continued investigation among PD researchers (Cobb, Zhao, and Dean, 2009; Feiman-Nemser, 2001; Seymour and Lehrer, 2006).

This gap between teachers' knowledge and practice across settings points to the importance of two critical issues: teachers' individual instructional expertise, and their personal identification with the normative representation of instructional competence across settings

(Gresalfi & Cobb, 2011). For example, a teacher who identifies with reform-oriented forms of instructional practices in PD settings might feel pressure to comply with the normative institutional identity reified by other teachers and leaders in their school. However, in a school with a normative institutional identity that privileges classroom discourse and inquiry in mathematics, this same teacher might thrive and develop very sophisticated forms of ambitious practice over time (Gresalfi & Cobb, 2011). Conversely, a teacher working in a school where others value and encourage ambitious practice may comply or even begin to personally identify with ambitious goals for student learning, but may struggle to enact such practice due to his or her own limited content knowledge or a lack of reliable classroom management routines that facilitate productive group work and whole-class discussion. Trying to explain teachers' practices strictly in terms of their expertise or agency, or strictly in terms of the professional development they receive to improve their practice, may offer less generalizable findings than research which identifies relationships between both the individual teacher *and* their institutional conditions.

In order to understand how teachers' school and district settings contribute to their improvement in instructional practice over time, we need research which examines aspects of the individual teacher (e.g., instructional expertise and their identification with particular forms of practice) as they are situated in particular settings that offer more or less support and accountability for their development. Understanding how particular supports can shape teachers' vision of practice or views of student capabilities – or alternatively, how teachers who identify with different forms of practice may engage differentially with available supports – can inform the improvement of professional development research and design, particularly in districts serving large numbers of diverse teachers. In this analysis, I investigate the interplay between several factors which prior research has found to influence teachers' longitudinal development of

ambitious instruction – their individual instructional expertise, district- or school-based supports for their learning, the instructional expectations of their school leaders, and their developing personal teaching identities.

## **Research questions**

This study was guided by the following research questions:

1. How are school-level learning opportunities and instructional expectations related to the quality of teachers' instruction over time?
2. How do teachers' instructional expertise, and their identification with particular forms of instruction, mediate the influence of instructional supports and expectations on their development of ambitious practice?

## **Background & Motivation**

Research in teacher education and professional development has contributed both theory and practice in supporting teachers to consider, engage with, and potentially develop an ambitious vision of instruction and coordinated classroom practices. Designs such as teacher time-outs (Lewis, Gibbons, Hintz, and Kazemi, 2015), teaching rehearsals (Lampert et al., 2013), and the use of written, verbal, or video depictions of classroom events illustrate the key elements of practice based professional education (PBPE; Ball and Cohen, 1999). PBPE designs provide opportunities for teachers to see more sophisticated practice, interrogate or otherwise unpack the goals and consequences of particular instructional moves, try out new instructional moves in service of these new goals, and reflect on the consequences of those moves for their students' learning as they plan new lessons and units (Ball, Hill, and Bass, 2005; Elliott et al.,

2009; Jackson et al., 2014). In accounting for teachers' learning and development of practice in these PBPE settings, researchers have identified a number of forms of expertise which are integral to teachers' enactment of ambitious practices – e.g., pedagogical content knowledge (Ball et al., 2005), their vision of what high quality instruction looks like (Munter, 2014), and how they frame students' potential to engage in rigorous math (Jackson, Gibbons, and Dunlap, in press).

While effecting change in the classroom, and thus in student learning, is the ultimate goal among those supporting and studying teacher learning in professional development settings, teachers' classrooms do not exist in a vacuum. Teachers' development of practice, even when supported by thoughtfully-designed opportunities to learn, is significantly influenced by the institutional settings they work in – in negative as well as in positive ways (TNTP, 2015; Little, 1989; Joyce & Showers, 1981; Franke, Carpenter, Levi, & Fennema, 2001; Cobb, Zhao, & Dean; Hill, 2011 ).

For example, factors related to teachers' working environments and collegial relationships which have been shown to be influential on their classroom practice include -

- the degree to which they have access to cognitively demanding instructional materials to use with students (Stein, Grover, and Henningsen, 1996)
- the opportunities they have to get individual support for their own classroom practice from others with greater instructional expertise (Olson & Barrett, 2004; Schifter, 1998; Ackland, 1991; Campbell & Malkus, 2009)
- what teachers understand themselves to be accountable for based on the professional norms and practices of their school leaders and other teachers in their schools

(Barraugh, 2011; McGinnis et al., 2004; Rousseau & Powell, 2005; Son, 2008; Yunpeng et al., 2006; Manouchehri & Goodman, 1998; Philipp et al., 1994).

Though the institutional setting of teaching is often hidden in studies of teacher learning in PD, this obscures the reality of teachers' work: the practices and expectations of others in the school community constitute more than just a backdrop for teachers' practice. Rather, they can and frequently do influence how teachers frame the goals of teaching, the materials they use when planning for or reflecting on practice, and the degree to which they have and engage in opportunities to learn. While many have noted the importance of attending to the institutional settings of teachers' practice when supporting teachers' learning over time, there has been less research that examines how teachers' longitudinal development of practice is shaped by aspects of their school settings in the context of instructional improvement efforts at the scale of a large school district.

To date, the best example available of such research comes from the *Middle School Mathematics and the Institutional Setting of Teaching* (MIST) project, a longitudinal mixed-methods study of large urban school districts that seeks to understand what it takes to improve the quality of math instruction through coordinated supports for teacher, instructional coach, and school leader learning. In an analysis by Smith, Rosenquist, et al (2013), researchers use longitudinal, quantitative survey data from more than 120 teachers spread over four school districts to test hypotheses about the relationships between teachers' school and district supports and their instructional practice, as measured through coded classroom videos. This analysis did reveal significant differences in instructional improvement between districts – namely, that only one district saw significant improvement in the quality of instruction over time. Though the nature of the supports offered in that district were distinct from the others, the analysis team has

found no clear-cut relationships between any particular forms of support offered within that district and the instructional improvement observed over those years. These findings may indicate that the *nature* or quality of the implemented supports in large districts vary, even when the particular *forms* of support are consistent from school to school within each district. For example, though a district may place (or provide funding for) instructional coaches at each school, the expertise of those coaches may vary significantly from school to school. Even within a school, the nature of a coach's work with teachers may vary depending on how many teachers the coach is working with and what other duties the coach is assigned. This analysis by Smith et al (2013) highlights the importance of further investigation into this school-level variation in the *quality* of supports for teachers' learning – which is one focus of the present analysis.

In summary, prior research indicates that in order to support students' development conceptual understanding along with procedural fluency in mathematics, teachers should be providing students with frequent opportunities to make meaning of math through discussion, justification, and problem-solving. However, developing ambitious instructional practices is challenging, and teachers need well-designed opportunities to learn with extensive support from more expert others. Even still, the quality (and quantity) of the learning opportunities they are provided is not the only determining factor in whether or not teachers go on to enact more ambitious practice in their classrooms. Teachers' development of ambitious instructional practices may be constrained or enabled by their pedagogical content knowledge (Ball & Bass, 2000; Hill, Rowan, & Ball, 2005; Howe, 1999; Simon, 1994), their vision of instruction (Munter, 2014), their current instructional practices (Visnovska and Cobb, 2015), their personal identification with the vision of instruction underlying support efforts (Greslafi & Cobb, 2011), or by the students they teach and how they frame those students' capabilities in mathematics

(Jackson, Gibbons, & Dunlap, in press; Wilhelm, Munter, & Jackson, in press). Teachers' instructional decisions are also influenced by their school and district context, which includes the nature of instructional supports they are provided and their perceptions of their leaders' instructional expectations. Thus, teachers' local environments may be more or less supportive of their development of expertise and capacity to enact ambitious instruction in their classrooms.

This portrait highlights the complex nature of schools and districts as settings for teachers' learning and practice. This complexity led me to ask why certain teachers develop ambitious instructional practices over time in the context of district-wide instructional improvement efforts, while other teachers in the same districts – and even in the same schools – actually *decline* in the quality of their instruction during the same time period. My goal in the present study was to pursue this question by examining a select group of teachers, their changes in knowledge, vision, and practice over time, and the implementation of supports and instructional expectations in their schools.

### **Brief Study Overview**

This dissertation study entailed a qualitative study of a small number of cases of teachers whose classroom video data suggest that they made atypical changes to their instructional practices over three or more years of the MIST study – changes which either improved or reduced the levels of cognitive demand in their lessons. I conducted a comparative analysis of the individual and institutional aspects of these teachers' practices over time, including their expertise (i.e., their math knowledge for teaching, their vision of high quality instruction, and their views of their own students' mathematical capabilities), their personal identification with (or resistance to) the vision of instruction central to reform efforts, the district- and school-based

supports they were given to improve their instructional practices, and the nature of their school leaders' expectations for their instruction. The goal of this comparative analysis was to identify potential explanations for these teachers' atypical progress or decline over time by comparing these two groups with each other and drawing on prior research to assess whether the teachers' school contexts were likely supports for their learning. I then explored the potential mediating effect of teachers' expertise and identities on the impact of supports and instructional expectations. This involved analysis and comparison of a third group of teachers: teachers whose instructional quality (e.g., the cognitive demand of their lessons) was stable and procedurally-oriented over time, though they worked in the same schools as teachers who significantly improved. These cases were used to explore conjectures in how teachers' instructional expertise, and personal teaching identification mediated the impact of instructional supports on their development.

### **Definition of Key Terms**

Throughout this dissertation I use the following terms to refer to aspects of teachers' practice and aspects of the school setting: *instructional expertise*, *instructional quality*, *supports*, and *instructional leaders*.

By *instructional quality* I refer both to the nature of a teacher's classroom instruction and the extent to which their instruction provides students with opportunities develop both procedural knowledge and to develop proficiency in the mathematical practices described in the CCSS-M. In distinguishing between forms of instruction that are higher or lower quality, I draw on work by Stein, Grover, and Henningsen (1996) as well as Boston and Wolf (2006) whose definitions of academic rigor of tasks, task implementation, and class discussions in mathematics are useful for

assessing the quality of student learning opportunities during a lesson. Generally speaking, by instructional improvement I refer to changes to mainstream classroom instruction which result in more opportunities for students to persist in solving problems, reason mathematically, develop and analyze mathematical arguments, and engage in conceptual discourse about their ideas. Thus a teacher's longitudinal trajectory of practice may be characterized by instructional improvement; furthermore instructional leaders may prioritize instructional improvement (at the level of the entire department or school) as a key strategy for improving student achievement.

By *instructional expertise* I refer to the forms of expertise which feature prominently in a teacher's classroom instructional practice – e.g., their pedagogical content knowledge, instructional vision, and views on students' mathematical capabilities. When referring to particular measures of teachers' changing instructional expertise, an *increase* in instructional expertise refers to a teacher's development of expertise associated with more inquiry-oriented forms of instruction. A *decline* in instructional expertise refers to a decline in a teacher's quality of classroom instruction, which may also entail changes in their vision of high quality instruction or their views of students' capabilities. I return to the topic of teachers' instructional expertise in the first section of Chapter II.

By *supports* I refer to the enacted forms of district- or school-based supports which were intended to help teachers develop in knowledge, vision, and/or practice associated with mathematics instruction. In this analysis, the supports of interest are pull-out professional development, instructional coaching, and school-based teacher collaborative time. I purposefully distinguish between *designed* supports and *enacted* supports because supports are frequently not implemented as intended due to inadequate time, expertise, or tools needed for these supports to have their intended effect. For example, a school leader's *design* for school-based collaborative

time might be that teachers use this time to explore mathematics together and rehearse launching tasks in order to improve their instruction. However, in practice, teachers may actually use the time to map out which standards they will ‘cover’ on which days during the coming weeks because no one in the group has sufficient instructional expertise to facilitate in-depth discussions and rehearsals of rigorous tasks.

By *instructional leaders* I refer to building-level principals or assistant principals who serve a supportive and evaluative role over some or all teachers in the building and make leadership decisions about school-based supports. By *instructional expectations* of instructional leaders I refer to teachers’ perceptions of school leaders’ expectations for what should be happening in mathematics classrooms – e.g., what curriculum teachers should be using, or what general forms of practice they want teachers to be using regularly. While teachers’ perceptions of their leaders’ instructional expectations may differ from the expectations leaders describe in interviews, this analysis focuses chiefly on the former, since I was chiefly concerned with how teachers negotiated between what they understood themselves to be accountable for and what they did in their classrooms. I return to the topic of instructional leaders in the last section of Chapter II.

## Chapter II

### Review of the relevant literature

In this chapter, I provide a brief overview of the empirical research about how teachers' instructional expertise, professional development, school leaders, instructional coaches, and teacher collaborative time (TCT) can influence (support or curb) teachers' instructional improvement over time. In the present study, I sought to understand how various forms of support in teachers' school contexts may contribute to their development of *ambitious* practices over time. For this reason, the following overview of the literature primarily features studies which have explicitly articulated *ambitious* instructional practice as the intended outcome of designs for supports and tools.

#### **Prior research relating teachers' knowledge, vision, beliefs, and practices**

A number of past studies in teacher learning have attempted to disentangle the nature of individual teachers' expertise, and how these forms of knowledge, beliefs, and vision contribute to their classroom practice. Additionally, a number of studies have found relationships between the "instructional starting points" (Cobb, Jackson, and Dunlap, 2014), or documented practices teachers enact at the start of an intervention or PD design, and their capacity or likelihood to enact specific high-leverage practices with support. These aspects of teachers' individual expertise can serve as both the focus of inquiry into teacher learning, and indicators or outcomes measured in studies of PD effectiveness. Based on the extant literature which relates aspects of teachers' instructional expertise and productive supports for their learning, the present

dissertation attends to four particular measures of individual expertise: teachers' initial instructional practices, their mathematical knowledge for teaching, their visions of high quality math instruction, and their views of students' mathematical capabilities. Though these aspects of expertise are discussed here one at a time (as they are often measured), in the present study they are considered in relation to one another, and in relation to the nature of the school settings in which teachers work.

**Teachers' Current Practices.** A critical dimension of teachers' instructional expertise which is relevant for understanding their ongoing development is the nature of their practice at the start of an intervention or support. One of the challenges of supporting teachers to develop reform-oriented practices is that they often have neither experience with, nor local models of such practices (Ball, 1996), and may see little reason to alter their typical routines. Thus, attending to teachers' current practice, and goals for student work, can be a fundamental first step in developing a shared motivation for teachers to engage in new practices. Visnovska & Cobb (2015) detail a PD design experiment in which the research team had to modify their design when they found that their designs for PD supports did not address a problem of practice that was shared with participating teachers. Once they found a common problem of practice on which to base their designs for PD (in their case, motivating students), teachers were able to make sense of these practices in relation to what they valued – and in turn to confront some of those values in light of more ambitious goals.

Though the issue of teachers' prior practice is clearly critical in assessing the appropriateness of the supports they are given for improvement, the nature of the data used in the present study precluded making such assessments. For example, in order to assess whether and

why a series of PD sessions about eliciting students' mathematical ideas during class discussions may have been critical in a teacher's improvement over time, it would be important to consider whether the teachers were already attempting to facilitate mathematical discussions, and whether the PD sessions were appropriate for building on their then-current practices. However, given the timing of data collection in the MIST project on which this analysis draws (e.g., the study began in the first year of these districts' reform efforts), there was insufficient data to draw on to make valid assessments of teachers' typical uses of specific practices in their instruction *before* the districts' reform efforts began.

**Mathematical Knowledge for Teaching.** Understanding and measuring teachers' mathematical knowledge for teaching (MKT) has long been the focus of inquiry into teacher characteristics that impact instruction (Ball and Bass, 2000; Hill, Rowan, and Ball, 2005; Howe, 1999; Simon, 1994). Early work by Simon (1994) and Howe (1999) clarified the distinction between mathematical knowledge and MKT, arguing that MKT is marked by knowledge about content in relation to student thinking and instruction. Hill, Rowan, and Ball (2005) built on this framework by developing an assessment for reliably measuring teachers' MKT in particular domains. They found that teachers with greater MKT were significantly more likely to see gains in their students' achievement scores in first and third grades. With regard to teachers' likelihood to enact *ambitious* instructional practices, Wilhelm (2014) found a significant relationship between teachers' MKT, their vision of high-quality instruction, and their likelihood to select and enact cognitively demanding tasks with students. Taken together, these findings suggest that teachers' MKT is relevant to their enactment of ambitious instruction, though I could not find any research that examined the relationship between teachers' MKT and their likelihood to

*develop* ambitious practice. In other words, it remains an open question whether teachers with more advanced MKT are more likely to develop ambitious practice in supportive contexts.

**Vision of High Quality Math Instruction.** Another component of instructional expertise which has been the subject of empirical investigation is teachers' *Vision of High Quality Math Instruction* (VHQMI, Munter, 2014; Wilhelm, 2014). This research has demonstrated that teachers' conceptions of what counts as good instruction, and the degree to which these conceptions are articulated in terms of student learning, are significantly related to teachers' likelihood to enact cognitively demanding mathematics tasks (Wilhelm, 2014). Further, a recent analysis by Munter and Correnti, (2014) found that teachers who had a more sophisticated, inquiry-oriented VHQMI at the outset of the MIST study were more likely to improve their practice over time.

These findings echo prior work by Thompson et al., (2013) that identified distinctions between the visions of practice of novice teachers who engaged in more or less ambitious instruction in science during their first years of teaching. They found that the novice teachers who had developed strong visions of practice that centered on student thinking went on to integrate a number of ambitious practices into the standard, district-issued curriculum they were given when they started teaching. They also found that these novice teachers identified more with the reform-oriented university community than with the school communities in which they worked. By contrast, novice teachers whose visions of practice were weaker, or organized around broad generalizations rather than specifications for practice, were more likely to experience tension in negotiating their membership between the university and local school communities.

This tension resulted in these novices enacting ambitious practices to a limited or nominal degree, despite being in comparable school settings as their counterparts.

Finally, Dunlap, Hunter, Rosenquist, and Smith (revise and resubmit) found that teachers' self-reports of improvement in their practice following PD on ambitious math instruction had no relationship with observed changes in their instruction in classroom video *except* in the cases of teachers whose VHQMI was inquiry-oriented. This finding indicates that not only does a teacher's VHQMI relate to their likelihood to develop more inquiry-oriented practice (as Munter and Correnti found), but their VHQMI also serves as a way of framing and monitoring their own learning.

Taken together, these findings around teachers' conceptions of high-quality instruction indicate that a teacher's VHQMI orients them to make sense of and develop particular practices in supportive contexts, and thus should be considered in studies of teacher learning over time.

**Views of Students' Mathematical Capabilities.** A fourth relevant component of teachers' instructional expertise is the ways in which they frame their students' mathematical capabilities. In their comprehensive coding analysis of teacher participants in the MIST project, Jackson, Gibbons, and Dunlap (in press) distinguished between what they term *productive and unproductive framings* of student difficulties in mathematics. The authors distinguish between *diagnostic framings* (e.g., the causes and sources of student difficulties) and their associated *prognostic framings* – what teachers considered appropriate responses to those difficulties. For example, teachers who framed student difficulties productively attributed these difficulties to the lack of adequate opportunities' to learn (including their own inadequate instruction), rather than an inherent lack of ability. Using the results of this comprehensive coding analysis, Wilhelm,

Munter, & Jackson (in press) found that these framings were significantly related to teachers' practices in orchestrating whole-class discussions. Specifically, they found that teachers with productive prognostic framings of student difficulties – e.g., they described supports that were aimed at supporting students' participation in cognitively demanding math instruction – were more likely to engage students in a rigorous whole-class discussion following student work time on a math task. Moreover, this relationship was more pronounced in classrooms with a high proportion of ELLs and students of color. This finding suggests that teachers of diverse populations who consider student capabilities to be a product of quality learning opportunities and productive supports are more likely to provide equitable access to mathematics through classroom discussions.

Given the problematic but pervasive assumption among many educators that students who lack basic skills or procedural competence cannot meaningfully engage in cognitively demanding problem-solving tasks, in the present study I specifically attended to teachers' ongoing framings of students' capabilities (their VSMC).

**Summary.** Taken together, these findings suggest that teachers' current forms of instruction are likely to influence their capacity or likelihood to enact certain high-leverage practices in the future, and that their development also depends on their individual knowledge, vision, and beliefs about their students. In the present study, in order to adequately understand improvements or declines in the quality of teachers' instruction over time, it was critical to understand these teachers' instructional expertise as they engaged with particular instructional supports.

## **Prior research on the influence of professional development on teachers' practices**

Given the critical role that teacher professional development is expected to play in most instructional reform agendas, there has been a growing attention to empirical research into 'what works' in teacher PD. This literature base includes large-scale statistical studies relating teachers' engagement in particular PD activities with their self-reported change in instruction (Garet, Porter, Desimone, Birman, & Yoon, 2001) as well as close studies of one or two teachers' experiences in PD (Jenlink and Kinnucan-Welsch, 2001; Naik & Ball, 2014; Sassi, Morse, & Goldsmith, 1997). Across these various research designs, some common features of effective PD have emerged. Jackson, Cobb, et al (2014) describe five core principles of high-quality teacher PD emerging from this body of literature, as well as from research on teacher education:

1. Supports for teacher learning should be sustained over time and involve the same group of teachers working together
2. Supports for teachers' learning should focus on issues central to instruction, organized around the materials that teachers use in their classroom
3. Supports should include co-participation with accomplished others in activities that approximate the targeted practices
4. Supports should include pedagogies of investigation and enactment – that is, activity structures which focus on the logic behind particular practices and their impact on student learning, as well as activities which support teachers to practice particular moves with the support of a more expert other
5. Supports should include facilitators who press on teachers' ideas differentially and build on their contributions

(adapted from Jackson et al, 2014, pp. 4-5)

In addition to these five principles, recent research from researchers investigating ways to improve teacher preparation through supported engagements with core practices suggests a sixth principle:

6. Supports should focus on supporting teachers to learn high-leverage practices – that is, research-based practices that have significant payoff in supporting student learning, that novices can feasibly learn relatively given appropriate support, and that they can incorporate into daily instruction (Grossman, Hammerness, & McDonald, 2009; McDonald, Kazemi, and Kavanagh, 2013).

One example of a high leverage practice is eliciting students' thinking (McDonald et al., 2013). As McDonald et al (2013) describe, eliciting student thinking involves “drawing out students' ideas about content and responding to those ideas in ways that move students' learning forward” (p. 382). For example, within an instructional activity such as a whole-class activity about a string of related problems (Kelley-Peterson, 2010; Parrish, 2011), teachers might elicit student thinking by them to share their strategies for solving a posed problem, or by pressing students to make connections between previously shared strategies. Scholars in teacher education at the University of Washington, Michigan State University, and the University of California, Los Angeles have developed routines for supporting pre-service teachers to develop high-leverage practices through cycles of investigation, rehearsal, enactment, and reflection (Kazemi, Franke, and Lampert, 2009; Kelley-Peterson, 2010; McDonald et al., 2013; M Windschitl et al., 2010; Mark Windschitl, Thompson, Braaten, and Stroupe, 2012).

Current work by Elliott (2016) and others at Oregon State University offers a helpful illustration of the six principles listed above. In their current research with practicing secondary mathematics teachers, the members of Elliott's research team have been adapting the learning

cycles developed by McDonald et al for use in PD. Their cycles also focus on supporting teachers to develop high-leverage practices embedded in instructional activities for high school mathematics classrooms. In a cycle, the ‘host teacher’ first chooses a focal high-leverage practice she wants to work on, such as eliciting student thinking (principle 6). The teacher then co-develops a lesson plan with the support of the research team and other teachers; whether this lesson is adapted from their own school’s curricula, or developed from scratch, it includes an explicit focus on repeatedly using the focal high-leverage practice (principles 2 and 3). The teacher then rehearses segments of the lesson with other teachers, using the rehearsal as an opportunity to refine or reformulate particular questions or activities in light of potential student responses (principles 3 and 4). As the teachers work together, members of the research team press teachers to share and explain their thinking about the math content, student learning, or teaching moves (principle 5). The entire group of teachers then separately enacts the lesson with their own students; afterward, the group meets again to collectively reflect on the lesson, revise the lesson plan, and sketch ideas for how to build on it in subsequent lessons (principles 2 and 4). The goal of this work is to support teachers to integrate these high-leverage practices into their daily instruction through repeated practice, focus, and attention to the resultant student learning.

**Teacher agency and identifying with reform efforts.** This research on PD reviewed thus far frames the challenge of instructional improvement as a problem of design and support for teacher learning. Within this framing, teachers’ lack of development of more ambitious practices would be explained by insufficient or perhaps developmentally inappropriate supports for their learning, and would signal the need for improvement of those supports. This framing positions the researcher or PD designer to more productively approach the problem of ineffective

supports than if the problem were instead framed as one of teachers' compliance or resistance. However, this framing does not necessarily preclude teachers' agency in mediating their engagement with available supports and, potentially, improved practices. Understanding the impact of PD and other supports for teacher learning on teachers' classroom instruction involves attending to not only the design of those supports, but also the extent to which teachers come to personally identify with the vision of instruction that underlies those supports – e.g., whether teachers come to adopt an inquiry-oriented vision of high quality math instruction, and hold themselves accountable for realizing such a vision in their own classrooms.

In the present study, I investigated the quality of pull-out PD and the potential relationships between this PD, teachers' personal identification with the goals of the PD, and their instructional improvement over time. This involved analyzing teachers' reports about the PD that they attended, the focus of this PD, and the kinds of activities teachers reported doing. In the case of PD sessions led by district or school level mathematics coaches, I also drew on data regarding these facilitators' instructional expertise to assess the likelihood that the PD sessions that case study teachers attended may have contributed to their improvement or decline over time. I also drew on teachers' interview data to determine how teachers position themselves with regard to their districts' instructional improvement initiatives (in particular, how they regard their districts' universal use of an inquiry-oriented math curriculum) in order to assess the extent to which they came to personally identify with the goals of the supports they were provided.

### **Prior research on the influence of instructional coaching on teachers' practices**

Studies of the efficacy of instructional coaching interventions date back to the 1980s, when 'peer coaching' emerged as a way to bolster the impact of PD workshops. Early reports by

Joyce and Showers, (1982) described how instructional coaches “used training procedures that virtually guarantee[d] the successful implementation of almost any approach” (p. 5). Studies of instructional coaching since this early work have offered more tempered recommendations for coaching practice, based on limited or mixed findings of the influence that instructional coaches can have on teachers’ practice (Neumerski, 2013). A number of models of mathematics coaching (or instructional coaching more broadly) draw on research-based activities from PD or teacher education in which teachers are supported to learn practice through modeling, approximations, and feedback from someone with more expertise (Gibbons, 2012; Grossman et al., 2009; Windschitl et al., 2012).

In their case study, Gibbons and Cobb (in press) studied one mathematics coach working one-on-one with teachers in their classrooms to support their development of ambitious practice in the context of district-wide instructional improvement efforts. They found that in addition to articulating a sophisticated vision of high quality math instruction, this coach constructed goals and learning trajectories for the teachers with whom she worked, and then supported the teachers’ development through modeling, co-teaching, debriefing, and co-planning. Gibbons and Cobb’s study is helpful in that this portrait of the case study coach offers insights into how a coach’s goals, knowledge, and practice were interdependent. However, given the design of this study, the researchers could not make direct causal links to instructional improvement in the coach’s school.

In a larger-scale study of a district-wide instructional coaching initiative with similar goals for coaches’ practices, Ai and Rivera, (2003) found that there were no significant differences between the practices of teachers who had engaged in modeling/observation and feedback activities with instructional coaches and those who had not. However, they also found

that the coaches in their study reported significant ‘resistance’ from teachers, and that not all the math teachers in their study felt that the principal was strongly supportive of the coach. These findings echo the findings of Matsumura, Sartoris, Bickel, and Garnier, (2009) who found that principals can crucially influence the extent to which coaches can work with teachers effectively. Moreover, the data from Ai and Rivera’s study were collected after the first year of the instructional coaching program. The authors conclude that their findings may also indicate that math coaches need more time (e.g., more than one year, and more time with fewer teachers) to have an effect on teachers’ practices. Though these problematic aspects of the study’s findings make it difficult to draw conclusions about the effectiveness of instructional coaching in supporting teachers’ development, it remains one of the few large-scale studies of coaching in the context of ambitious, district-wide goals for teachers’ learning.

In a robust randomized control study of coaching practice, Campbell and Malkus (2009) used daily activity logs with 24 mathematics coaches over several years to understand the kinds of activities that they were doing with teachers. These math coaches had received substantial professional development before beginning their work with teachers, including support to develop their math subject matter knowledge especially. In analyzing changes in the coaches’ daily logs over time, the researchers found that over the three year period of data collection, the coaches in their study spent less time on one-on-one coaching, and spent more time doing administrative tasks, communication and materials coordination, and assessment-related duties. Using state-provided student assessment data, the researchers then used hierarchical linear regression to determine whether these coaches’ work had a measurable effect on student achievement in the schools when compared with control schools in the same district. They found that there was a measurable effect on student achievement in schools that had math coaches,

though the effects were not statistically significant until the coaches had been working with the same teachers for more than 1 year. One limitation of their analysis is that because data about the coaches' practices came from the coaches' daily logs (rather than from individual teachers), the data could not be used to investigate relationships between specific coaching practices and student achievement. Notwithstanding this limitation, Campbell and Malkus's analysis provides important evidence that math coaching can be a productive school-level strategy for improving student achievement through supporting teachers' learning.

These findings provide a number of testable conjectures about how instructional coaching can support improved student learning outcomes by providing opportunities for individual teachers to develop their practices. Gibbons and Cobb's (in press) findings suggest that an instructional coach's capacity to support teacher learning in one-on-one coaching is likely related to her own expertise in enacting more ambitious instruction, and the kinds of activities she engages in with teachers (such as modeling, co-teaching, and doing mathematics with teachers one-on-one or in groups). Ai and Rivera's (2003) findings suggest that instructional coaches also need to build trusting relationships with teachers, which may be related to the ways they are positioned by their school leaders and subsequently perceived by their teachers (Matsamura et al, 2009).

In the present study, I investigated these research-based conjectures about instructional coaching as a support for teacher learning by analyzing transcripts from interviews with teachers, instructional coaches, and instructional leaders in order to:

- determine the instructional expertise of any instructional coaches that the case teachers worked with (e.g., whether they had a math teaching background, what their vision of high quality math instruction was)

- determine the nature and quality of support that the case teachers received from their coaches, including the kinds of activities they did together
- investigate whether the case teachers attended PD or TCT that was facilitated or attended by a math coach, and if so, what the focus and central activities of those meetings were (i.e., whether these meetings involved high-leverage facilitation moves, such as the coach leading the teachers in an investigation of math)

### **Prior research on the influence of teacher collaboration on teachers' practices**

A number of studies have documented how teachers who work together closely in regular collaborative meetings can develop shared epistemic stances, shared language for talking about or categorizing students, and a shared pool of materials for practice (Hall and Horn, 2012; Horn, 2007; Lampert, Boerst, and Graziani, 2011; Little, 2002). For example, Horn (2007) illustrated through a comparative case study how teachers' labels for students based on their achievement became taken-as-shared among teachers in a collaborative workgroup, and resulted in inequitable groupings of students in mathematics. This problem of language, categorization, and instructional inequity is especially visible when accountability test data is used as the primary indicator of student learning. Horn, Kane, and Wilson (2015) report on two groups of teachers who brought very different meanings and interpretations to data from the same district assessment. The authors illustrate how the different epistemic stances which became shared among teachers in the two workgroups influenced how teachers were positioned to learn about practice from these conversations. In a more positive example, Lampert and Graziani's (2011) description of an Italian Language School provides a model case of a teaching community whose shared tools of teaching (e.g., lesson plans with embedded pedagogical reasoning) and

collaborative planning and debriefing routines contribute to the community's development of a shared vision of practice. Brasel, Garner, Horn, and Kane (2015) provide a framework for identifying the potential learning opportunities that surface in teachers' workgroup discussions of student achievement data in particular. They illustrate how teachers' opportunities to consider the interplay between students, teaching decisions, and mathematics rest on the extent to which they discuss:

- *What* do we need to reteach?
- *To whom* do we need to reteach it?
- *Why* did students struggle with this?
- *How* do we reteach it? (Brasel, Garner, Horn, & Kane, 2015, p. 1).

Discussions of the first two questions to the exclusion of the latter are very common in teacher workgroups, where teachers identify the standards and students with the poorest results and group them for remediation, either in re-teaching lessons during class or in secondary supports. More challenging is to unpack the gaps in student thinking and learning opportunities that contributed to their failure, and to problematize prior instruction accordingly (the *why* and *how*).

In my Major Area Paper (MAP; Dunlap, in preparation), I conducted a synthesis of the TCT literature and identified three critical components of ambitious TCT in which teachers developed more ambitious practices, and a number of school conditions which were necessary to support these components. The following paragraphs provide a summary of these findings.

The first component of productive TCT which I identified was that *time is organized around high-level activities*. Putting teachers together in a room to work together does not necessarily create meaningful opportunities to develop more ambitious goals and practices in

their instruction (Grossman, Wineburg, and Woolworth, 2001; Levine and Marcus, 2010; Sassi, Bopardikar, Kimball, and Michaels, 2013; van Es, 2011). Rather, as Levine and Marcus, (2010) point out:

Decisions about the structure and focus of teachers' collaborative activities can both facilitate and constrain what teachers can learn together by influencing: whether teachers make their own practices in the classroom public; which aspects of teaching are discussed; the degree of specificity with which teachers share aspects of their work; and the kinds of information about students teachers make available to each other (p. 397).

Levine and Marcus echo others' findings that the degree to which activities connect representations of teachers' classroom practice and their knowledge of teaching, content, and student learning has considerable impact on the nature of the learning opportunities TCT offers to teachers (Crockett, 2002; Horn, Kane, and Wilson, 2015; Little, 2012). To this end, the empirical accounts of productive TCT I reviewed featured a number of TCT activity designs principled on this attention to *making practice public* through shared representations, *developing deeper content knowledge through collaborative exploration*, and *developing deeper understandings of student thinking* through the examination of student explanations and work samples. Looking across the empirical accounts of productive TCT, all of the articles which detailed teachers' activities in TCT incorporated at least two of these three themes to some degree (*making practice public*, *developing deeper content knowledge*, and *developing deeper understandings of student thinking*). However, as nearly all of the authors noted emphatically, the activities themselves don't cause teachers to learn. While the *designs* of these TCT activities were high-level in that they were informed by theories of how teachers learn new practices in

collaborative settings, the *implementation* of these activities was productive due to a number of enabling conditions in the TCT settings as well as in the larger school and district contexts.

The second component I identified in my MAP was that productive TCT was facilitated by someone with relatively more expertise in ambitious instruction. Across the reviewed studies, facilitators generally served three main purposes: supporting teachers to learn to identify critical moments of student thinking or pedagogy (Stevens & Hall, 1998), pressing teachers to make explicit connections between pedagogical choices, instructional content, and student learning (Little, 2012), and supporting the development of collaborative norms for joint work among teachers in the group (Wenger, 1998).

The third component was that TCT was just one part of a larger approach to improving classroom practice through supporting teacher learning in various professional development settings. For example, in a number of the studies, teachers also attended more intensive summer or weekend PD workshops facilitated by the research team or PD organization (Cobb, McClain, de Silva Lamberg, and Dean, 2003; Grossman et al., 2001; Kazemi and Franke, 2004; Wood, 2007; Zech, Gause-Vega, Bray, Secules, and Goldman, 2000). Some of the studies also provided teachers with one-on-one coaching or classroom observations and feedback outside TCT (Ermeling, 2010; Hindin, Morocco, Mott, and Aguilar, 2007; Zech et al., 2000). Additionally, as mentioned previously, teachers in a number of studies were given new curriculum materials to use to support their development of more ambitious practices (such as the CGI math tasks in Kazemi & Franke, 2004). These studies illustrate part of what Bryk, (2010) termed a “coherent instructional guidance system,” which articulates the “what and the how” of instruction through curricular materials, assessments, and professional development (p. 24). Bryk found that schools that improved student learning outcomes “used professional development as a key instrument for

change... and when teaching was guided by a common, coherent, and aligned instructional system” (p. 26). This suggests that districts implementing TCT as a support for teacher learning will be more effective if they focus on developing the same instructional vision across other “learning events” (Cobb and Jackson, 2012) offered to teachers.

The small research base suggests that if TCT is organized around high-level activities, led by a relatively expert facilitator, and is part of a coherent instructional system for supporting teacher learning, TCT can be a meaningful support for teachers to develop more ambitious practices. Using teacher and coach interview data, I explored these possible relations in the present study by:

- investigating whether a member the case teachers’ TCT group (especially the facilitator, if there is one) had significant expertise in ambitious instruction
- investigating the forms and foci of activities that case teachers reported doing with others in TCT, and whether these activities, coupled with the available expertise in the group, were likely to support the case teacher to develop instructional expertise
- investigating whether the focus of activities in TCT in case teachers’ schools was consistent (or potentially in conflict) with the instructional goals orienting other supports for teachers’ learning, such as district PD or one-on-one coaching

### **Prior research on the influence of principal leadership and expectations on teachers’ practices**

In drawing together findings about the role of leadership in teachers’ learning and development, I found the reviews by Neumerski, (2013) and Larbi-Cherif (2015) to be especially helpful. Neumerski’s review offers a historical and thematic synthesis of findings about

instructional leadership, including principal, teacher, and coach leadership in schools. Larbi-Cherif's review focuses specifically on leadership functions and tasks which may improve the quality of instruction in schools. Both of these reviews of the literature indicate that there is very little research available that connects specific principal practices, or their reasoning behind these practices, to teachers' development of ambitious practice (Neumerski, 2013; Larbi-Cherif, 2015). Furthermore, they both conclude from their reviews that effective principal leadership tends to be more about principals' work in creating school environments which support teachers' learning than about their direct work with teachers (e.g., through classroom observation and feedback).

Such indirect leadership practices include supporting ongoing teacher learning through the provision of professional development and instructional coaching, and creating time in the school day for teacher collaboration (Larbi-Cherif, 2015). The findings of Rigby et al (revise and resubmit) add additional evidence that indirect leadership practices may be a more worthwhile than direct, one-on-one leadership practices. They found that though many school leaders are expected to conduct regular classroom observations followed with feedback to teachers, most school leaders lack the capacity or content expertise to provide the kinds of feedback to teachers that would support their development of more ambitious practices.

Specific to creating conditions for productive teacher collaboration (TCT), my synthesis of the TCT literature (Dunlap, in preparation) indicates that principals can potentially create these supportive conditions by:

- Scheduling time during the school day for TCT
- Arranging teaching teams with consideration for levels of expertise represented in each team
- Providing teachers with high-quality curricular materials

- Identifying TCT facilitators based on their relative instructional expertise, (instructional coaches) and positioning them as a support for teachers' learning
- Providing support for facilitators to improve their own practices
- Communicating clear goals and expectations for TCT
- Monitoring and pressing for improved instruction
- Protecting TCT from competing initiatives that would impinge on teachers' opportunities to learn in TCT

In sum, while there is very little empirical research which has examined the role of school leaders in improving the quality of classroom instruction, this literature has put forth some testable conjectures: teachers are more likely to develop ambitious practices when they and their school leaders share a common and ambitious vision of high-quality instruction (Nelson, 1997; McGhee and Lew, 2007; Larbi-Cherif, 2015), when teachers understand themselves to be accountable for enacting ambitious practices in their own classrooms (Leithwood and Riehl, 2003), and when their leaders have organized ongoing supports for them to improve their practices with more expert others (Spillane, Diamond, Walker, Halverson, and Jita, 2001; Printy, 2008).

In the present study, I drew on prior analyses of principals' vision of instruction, along with teachers' interview data in order to understand *how school leaders' practices in the case teachers' schools may have contributed to the improvement or decline in the quality of teachers' instruction over time*. Given the findings of Larbi-Cherif's (2015) review of the literature which suggest that school leaders may have more impact on teacher learning through *indirectly* supporting their learning, I looked for evidence of these indirect forms of support in my analysis.

This involved

- investigating the case teacher's perceptions of his/her school leaders' (principal and/or AP) expectations about their classroom teaching and their likelihood to influence teachers' practice in particular ways
- investigating the case teacher's perceptions of his/her school leader's expectations for what teachers do during TCT.

## **Summary**

The research literature reviewed in this chapter indicates that in order to make sense of teacher learning in relation to their school contexts, it is critical to take into account teachers' instructional expertise, their personal teaching identity, what they understand themselves to be accountable for, and the supports they were provided (and engaged with) to develop more ambitious practices.

While some of these aspects of teachers' school settings have received less attention or empirical research which examines their impact on teachers' development, there are even fewer examples of research that examines more than one of these aspects at the same time. For example, in my review of the literature on TCT (Dunlap, in preparation), I found no examples of empirically-grounded recommendations for the role of principals in creating school-based TCT that supports teachers' development of ambitious practice. There has been marginally more research that looks at the intersection of instructional coaching practice and instructional leadership practices (Matsumura et al, 2009; Mangin, 2007; Wren & Vallejo, 2009). However, most of this research stops short of tracing the effects of these coaches' and leaders' practices to their impact on teachers' knowledge, vision, or practice. By contrast, a central goal of this

dissertation study is to bring together conjectures from the literature on teacher expertise, instructional leadership, and particular supports for teacher learning, and explore these conjectures in longitudinal cases of pronounced change.

## Chapter III

### Methodology

#### Research Design

The goal of this study was to better understand how instructional expectations and supports for teachers' learning are related to teachers' development of ambitious practice within school districts pursuing instructional improvement at scale, and how these relationships are mediated by teachers' current practices, knowledge, and teaching identities.

I return now to the research questions for this study to orient the research design I used to address them.

1. How are school-level learning opportunities and instructional expectations related to the quality of teachers' instruction over time?
2. How do teachers' instructional expertise and their identification with particular forms of instruction mediate the influence of instructional supports and expectations on their development of ambitious practice?

Assessing the degree to which the *nature* of these expectations and the *quality* of these supports may have contributed to teachers' instructional improvement – i.e., getting to the “how” – implicates the use of a qualitative study of strategically selected cases. In order to investigate the potential contribution of the supports outlined above to teachers' development, I drew on data from the Middle School Mathematics and the Institutional Settings of Teaching (MIST) study, an organizational design study of four large urban school districts attempting to improve the quality of middle school math instruction. These districts were selected for study due to the ambitious

nature of the instructional visions their district leaders were pursuing, and the forms of support being provided to teachers. The MIST project involved collection of numerous forms of data from teachers, math coaches, school leaders, and district leaders in order to gauge the quality of enactment of district-level and school-level strategies for supporting teachers to develop inquiry-oriented practices. For the present analysis, I drew on data from the two districts in the MIST project whose district strategies included providing middle-school math teachers with the supports of interest: instructional coaching, math-specific professional development aligned with the rigorous, inquiry-oriented curriculum adopted by the districts (*Connected Mathematics Projects 2, CMP2*), and school-based TCT - though the latter was implemented in varying ways throughout these districts. A description of these districts (Districts B and D) follows in the next section.

Teachers in the other two MIST districts (Districts A and C) were excluded from the present sample for several reasons. First, district C was excluded because the quality of curricular materials being provided by the district (and thus being used in most classrooms) was substantially lower than in the other three districts. Relevant to this analysis, this difference in district C poses challenges to making comparisons among teachers and their opportunities to learn. The challenge of teaching cognitively demanding lessons in this district would involve not only *implementing* rigorous curricula, but also finding or creating that curricula to begin with. District A teachers were excluded from the sampling pool because many of the schools in District A were so small that there was very little math-specific TCT across the district. Furthermore, neither district A nor C included designs for instructional coaching in their district-wide strategies for instructional improvement.

**Addressing RQ 1.** Flyvbjerg (2006) points out that the in-depth study of a small number of extreme cases can illuminate the mechanisms behind particular phenomena in ways that representative cases cannot. Subsequently, these mechanisms and their explanatory power can then be investigated in other samples. Addressing my first research question involved comparative analysis of two groups of cases selected based on their extreme, or atypical, nature. Given my interest in testing and generating conjectures about the ‘mechanisms’ of support for teachers’ learning, the first set of cases was comprised of teachers whose instructional quality atypically improved over multiple years in order to generate and test conjectures about critical features of these teachers’ school and district settings. I selected four teachers who showed atypical improvement over three or more years: two in District B, and two in District D. I intentionally selected two atypical improvers in each district in order to look for similarities between their experiences in pull-out PD, much of which was offered by the district. In addition to selecting atypical cases of teachers whose instructional quality improved over time, I also analyzed four contrasting cases of teachers whose instructional quality rapidly and significantly declined over time. The goal of studying these ‘negative’ cases was to add robustness to the conjectures tested and generated in the study of positive cases, and to develop additional conjectures about the interrelationships between various aspects of the school setting that are relevant to teachers’ learning. The inclusion of these cases was meant to safeguard against oversimplifying the relationships between supports for teacher learning and teachers’ instructional improvement. The primary data for this comparative analysis of atypical improvers and decliners were transcripts from annual teacher interviews, in which case teachers were asked about the forms of support they were provided by their school or district for improving their instruction or implementing CMP2. (Further description of the data to be used for analysis begins on p. 59.)

**Addressing RQ2.** My second research question asks whether the influence of instructional supports and expectations may be dependent on an individual teacher's instructional expertise or the degree to which they personally identify with the goals of the support strategies. For example, instructional supports may have very little influence on a teacher's practice if she doesn't view all of her students as capable of engaging in cognitively demanding tasks – even if her vision of high quality instruction involves such tasks. Alternatively, a teacher who does not personally identify with the vision of instruction constructed in PD, coaching interactions, or TCT may resist this identity (e.g., speak out against what is being advocated in PD) or even comply with it in those settings with no intention of making changes to her instruction (Gresalfi & Cobb, 2011). In the former case, the teacher's instructional expertise may limit her development of particular practices. In the latter case, a teacher who has not come to personally identify with the goals of the reform efforts will likely lack sufficient motivation to meet the challenges of ambitious instruction – especially if there is little accountability for doing so.

To explore these contingencies, I also analyzed a third group of cases: teachers whose instruction remained stable and procedurally-oriented though they worked in the same schools (during the same years) as the case teachers who showed atypical improvement. Introducing these cases allowed me to investigate how a teacher's expertise and teaching identity can mediate the influence of instructional supports for teachers within the same school. For all three groups, I drew on case teachers' instructional expertise data to determine whether there are categorical differences between measures of their knowledge, vision, views of student capabilities, and practice that may relate to their differential trajectories in practice over time. To assess the extent to which teachers personally accepted, complied with, or resisted the goals of their districts' reform efforts over time, I drew on segments of interviews with these teachers in which they

were asked whether they considered CMP2 appropriate for all students, and the extent to which they felt others in the school shared their vision of high quality instruction.

Table 1 below summarizes these research design decisions and their connection to my research questions and the broader phenomena my cases were chosen to represent.

Table 1		
<i>Summary of Research Design</i>		
Research Question	Cases for analysis	What will these teachers constitute cases of?
How are school-level learning opportunities and instructional expectations related to the quality of teachers' instruction over time?	Teachers in the MIST project whose instructional quality atypically improved (4) or declined (4) over 3 or more consecutive years.	Teachers in schools whose districts are pursuing instructional improvement in mathematics through coordinated supports for teacher learning (specifically coaching, TCT, and pull-out PD) and instructional leadership.
How does teachers' instructional expertise and knowledge mediate the influence of instructional supports and expectations on their development of ambitious practice?	Teachers whose instruction remained fairly procedurally oriented, despite working in the same school (during the same years) as those who atypically improved.	Same as above, with additional attention given to their instructional expertise in relation to the focus of supports and instructional improvement strategies. For example, a teacher may be a case of teachers with unproductive views of student capabilities whose school supports implicitly assume but are not designed to specifically develop more productive views of student capabilities.

### **Research setting: Districts B and D**

This study included schools in two different districts, located in two different southern US states (District B in State B, and District D in State D). This analysis focused primarily on the school settings in which case teachers worked. However, the school-level policies, instructional expectations, and supports for teacher learning central to my analysis were created, informed, or provided by district leaders in Districts B and D. Thus, understanding these teachers'

development within their school contexts also requires a broader understanding of the state and district contexts surrounding their work.

During the years of the MIST project's data collection (2008 – 2015), these districts implemented (and subsequently revised) a number of strategies related to teacher support and accountability. In the next few sections, I provide a brief description of these district contexts, the district-level supports and expectations for middle-grades math instruction, and changes to state policy and district leadership that impacted these supports and expectations over the years of the larger study.

**District Vision.** The MIST districts were not immune to the challenges facing most urban districts across the US. For example, both districts served a large proportion of students from low-income households, and District B in particular served a large percentage of English Language Learners (ELLs). Both districts also faced challenges associated with teacher and school leader turnover and student mobility. Furthermore, both districts had a number of schools under threat of state sanctions due to chronic low performance on state assessments. Despite these common challenges, what made these districts atypical was that their district leaders articulated ambitious goals and strategies for raising student achievement in mathematics by improving the quality of mainstream classroom instruction. From 2008 to 2012, or years 1-5 of the MIST study, Districts B and D implemented several district-wide strategies designed to provide teachers with support opportunities to improve their instructional quality. Guiding these strategies were leaders in both districts who outlined particularly ambitious goals for classroom instruction in middle schools. For example, the superintendent of District D led his district leadership team in drafting a “Vision Document” intended to orient their reform efforts. This

document described instruction that emphasized students' development of both conceptual understanding and procedural skills as they engaged in cognitively demanding tasks and rich classroom discourse. Similarly, the district math leaders in District B also articulated an ambitious vision of math instruction that oriented their improvement strategies.

### **Key District Strategies.**

**Curriculum and PD.** Based on these ambitious goals for teachers' practices, both districts adopted the CMP2 instructional materials for their middle school math programs – an inquiry-oriented that which organizes units around thematic tasks in which students solve a small number of related, open-ended problems before discussing their ideas in a whole-class discussion facilitated by the teacher. These districts also created (and periodically revised) curriculum framework tools intended to support teachers in pacing their instruction within a timeline that included common assessments based on CMP2, as well as district benchmarks which approximated the state tests. During years 1-5, district math leaders in B and D also provided teachers with professional development intended to support them in developing practices central to effectively implementing CMP2 (e.g., launching tasks without reducing them to a set of procedures; using talk moves to elicit and connect student ideas in discussion).

**Coaching.** In addition to curricular tools and pull-out PD, both districts developed instructional coaching programs designed to support teachers in their own classrooms as they attempted to implement CMP2. In District B, math leaders acquired a state grant to support middle school students' development of the mathematics skills and understanding they would need to be successful in high school algebra. The district used this grant to provide every middle school with funding for a part-time mathematics instructional coach, who would be chosen or

hired by the school principal. This resulted in significant variation in the expertise and practices of math coaches across the district; many had no more instructional expertise for implementing rigorous problem-solving tasks than the teachers with whom they worked. District math leaders therefore provided pull-out professional development for the math coaches intended to improve their own teaching practices, as well as their capacity to support other teachers' learning.

District D's instructional coaching model included two types of coaches: district coaches who reported to the district math leader, and school-based coaches who reported to their principals. The district coaches divided their time between working on district assessments and curricular tools, and working with teachers in multiple schools. Prior analyses of these district math coaches' prior teaching practice, their visions of instruction, and their math knowledge for teaching indicated that they had significantly more instructional expertise than the majority of middle-grades math teachers in the district sample. However, they each served several schools and were therefore limited in the time they could spend with teachers from any given school; thus, the coaches had to prioritize which classrooms they visited, and often did so based on student achievement scores. Additionally, some middle schools in District D also had full-time school-based staff developers (SDs) who were funded by an external corporate grant. These staff developers (SDs) were meant to support math *or* science teachers, depending on their own content backgrounds. MIST analyses indicated that like the district coaches, most math SDs also had more instructional expertise than the teachers they worked with. Similar to District B, District D math coaches were also supported by the district math leader with pull-out PD, though the focus and amount of these PD sessions varied by year.

***Teacher collaborative time.*** Both districts viewed school-based collaborative time as an opportunity for teachers to work together to prepare for and improve their instruction. School

leaders in both districts were encouraged to provide teachers with regularly-scheduled time to meet and plan for instruction, though the frequency and duration of these meetings was largely left to the principal to arrange. By year 5, all school leaders in District B were expected to schedule time each week for mathematics teachers to meet, preferably in grade-level teams. In District D, negotiations with the teacher's union prevented the district from changing district-wide policy regarding how often teachers met in collaborative time. However, by year 6 most schools in District D had regular (weekly or bi-weekly) TCT for math teachers, often in grade-level groups, incorporated into their schedules.

***School-based instructional leadership.*** In both districts B and D, district leaders expected school principals to oversee and effectively implement the above supports. In some years, principals were provided with content-general professional development to help them recognize high quality classroom discourse across content areas. In other years, the district provided math-specific PD to school leaders, including sessions co-designed and co-led by members of the MIST research team.

**Changes beginning in year 5.** In both districts B and D, there were important changes made to state and district policies that affected math teachers. At the state level, both districts' states adopted more rigorous mathematics standards in year 5. In many cases, content standards were moved into lower grades (e.g., more statistics and probability content standards appeared in elementary grades than before), while across grades students were expected to develop proficiency in a range of mathematical practices including forming mathematical arguments and critiquing the reasoning of others. In year 5 both states also adopted new assessments that were intended to assess these more rigorous standards. Many teachers reported finding difficulty to

reconcile the CMP2 instructional materials with the new standards – possibly because they were not explicitly organized around the new standards, and in some cases because some standards had moved into another grade level. The districts provided teachers with revised curriculum frameworks that reorganized the sequence and pacing of instruction to align with the new standards, and provided teachers with optional PD meant to familiarize them with the new standards. However, many teachers still found it difficult to find standards-aligned resources on a daily basis.

In year 6, reorganization and new leadership in District D brought a number of changes to the district’s vision of instruction and central strategies for improving student achievement. For example, the new leaders relaxed expectations that teachers should teach in any particular ways or use any specific curricular materials, and instead encouraged teachers to locate and use any materials they found which were aligned to the new standards. Many schools purchased new, procedurally-oriented textbooks (e.g., lessons featured worked examples followed by practice problems). Across the district, teachers worked individually or with their grade-level teams to locate various internet resources for day-to-day instruction. MIST analyses found a steady decrease beginning in year 6 in the rigor of mathematics tasks that teachers in the sample used with their students.

Year 6 also marked significant changes in both districts’ instructional coaching initiatives. The formerly district-based instructional coaches in district D were placed in schools, and all instructional coaches in the district (including SDs) were expected to provide support to all teachers in the building, regardless of content area. All coaches were also expected to help organize and TCT meetings across the school, and in some cases facilitate TCT meetings where the principal was concerned the teachers would not use the time effectively. However, outside of

a directive to use data to improve student achievement, the district provided little guidance for how TCT should be used, and provided minimal training to coaches to be effective facilitators.

In District B, district leaders decided in year 5 to dissolve the school-based coaching program after seeing no clear positive impacts of the program on student achievement or instructional improvement. Schools who wanted to continue to fund their math coaches had to locate their own funding to do so. Three math coaches and a senior math leader were retained at the district level, and were responsible for providing math-specific professional development to teachers in their schools. Although these math coaches and their supervisor were markedly more expert in ambitious math teaching than the teachers in their district, they were also given a large number of responsibilities, which resulted in limited support they could provide to individual teachers or TCT groups.

Finally, another shift in both districts between years 5 through 8 was increased focus on student achievement outcomes by all teachers, instructional coaches, and school leaders in both districts. In both districts, the increased focus on students' achievement scores was often palpable in the latter years of the study. For example, in District D, many schools created 'data rooms' in which teachers, coaches, and school leaders used large charts and spreadsheets to track individual students' progress through state and district benchmarks during the year; these rooms were sometimes jokingly called 'war rooms' by many principals for their resemblance to military strategy operation offices. As another example, limited resources for providing additional supports or tutoring to failing students led many school leaders to focus their efforts on "bubble kids" who were the closest to passing the state assessments (Booher-Jennings, 2005). These students were often targeted for 'triage' of second and even third math classes, after-school and

weekend tutoring, and other supports because they represented the potential for high return on instructional investments in terms of the school's rankings.

## **Participants**

Participants in the MIST study included teachers, instructional coaches, principals, assistant principals, and district leaders<sup>1</sup>. The latter group included members of all district central office units that worked to support the improvement of mathematics improvement through developing curriculum, PD, support for coaches or school leaders, or supplemental supports for math teachers working with ELLs, or students with exceptional needs.

In 2006-2007, MIST PIs worked with district leaders to select a sample of schools (6-10 each in districts B and D) that would represent a range of student demographics and achievement histories. In year 5 of the study (2011-2012 school year), 5-6 schools were added to the sample in each district, again selected in collaboration with district leaders so that the entire sample of schools in each district would be more or less representative of the district's middle schools. In both school recruitment years, MIST project PIs approached the selected schools' principals in order to ask for their school's participation.

In the final (eighth) year of the study, the number of schools in the MIST sample was reduced to a minimum amount required to provide the two districts feedback on their strategies,

---

<sup>1</sup> While this analysis will draw on prior summaries and syntheses of interview data from district leaders to provide contextual information about particular district strategies (e.g., district-wide expectations for coaches), those interviews will not be the focus of the present analysis, and thus are not discussed in depth here.

and a number of instruments and measures (including classroom video data collection and a number of interview questions) were dropped. For this reason, I did not draw on data from year 8 of the study for this analysis.

In all, the sample of schools in the MIST study from districts B and D from study years 1-7 includes:

- 11 schools who participated for 7 years
- two schools that participated for 6 years
- 10 that participated for 3 years
- one school that participated for 2 years
- two schools that participated for 1 year

The latter two groups of schools above were excluded from the sample since teachers in those schools participated in only one or two years of data collection.

Each fall, members of the research team visited schools to randomly select and then invite between four and seven math teachers to participate in the study. Preference was given to teachers who had previously participated in the study in order to generate a longitudinal dataset. Teachers who left the study or left the school from one year to the next were replaced by randomly selecting other teachers in the same schools and inviting them to participate. Teachers who primarily taught exceptional education classes, supplemental support math classes, or accelerated math classes were excluded, in order to restrict the sample to teachers who primarily taught mainstream math classes.

Along with teachers, any instructional coaches (math-specific or not) who worked regularly with participating mathematics teachers were also invited to participate in the study. Thus, our sample of coaches included math content coaches with substantial experience and

expertise in math, along with generalist coaches who did not have a math background but regularly attended math department or grade level math meetings at the request of the principal or district leaders.

School principals, along with any assistant principal who worked with mathematics teachers regularly were also invited to participate. Thus the sample of school leaders included assistant principals who were assigned to lead the math department specifically because of their content knowledge, for example, or because the principal had divided up the faculty evenly across instructional leaders for teacher evaluations or observations.

## **Data and Measures**

In this section, I describe the principle data sources that were used in my analysis, the timing and procedures for their collection, and their relevance to my research questions.

All school-based participants in the MIST study completed two forms of data collection each spring: a 45 minute to hour-long audio-recorded interview, and an electronic survey. In addition, math-specific coaches and teachers also took a paper-and-pencil assessment of their math knowledge for teaching, the *Learning Mathematics for Teaching* (LMT) assessment, developed by Hill, Ball, and Schilling, (2008). Finally, the project also collected classroom video data from each participating teacher or part-time coach who also taught a mainstream math class. For each teacher in each year, we recorded two consecutive lessons in the same class period.

Collection of these data occurred each spring, when members of the research team collected data in four waves: interviews (January), classroom video data (typically February through March), electronic surveys (February-March), and paper-and-pencil tests of teachers' mathematical content knowledge (March).

**Assessing teachers' instructional expertise and identification.** The present analysis drew on four primary measures of teachers' instructional expertise: an assessment of their instructional quality derived from coding their classroom video, a paper-and-pencil assessment of their math knowledge for teaching, and interview-based measures of their vision of high-quality math instruction and their views of students' mathematical capabilities. Though selection of teachers for this analysis used measures of teachers' instructional practice (the IQA, described below), the other measures of their instructional expertise described in this section were used to draw inferences about the relationship between teachers' expertise and the influence of the various supports or instructional expectations in their school and district settings.

*The Instructional Quality Assessment.* Each year, each participating teacher worked with the MIST project manager to schedule two consecutive days of instruction to be videotaped by a project team member during the same class period. The goal of taping two consecutive class days was to capture at least one entire problem-solving lesson, which could extend over two days to allow for longer class discussions. Consistent with prior uses of the Instructional Quality Assessment protocol (IQA, Boston and Wolf, 2006), teachers were told in advance that we were hoping to film a problem-solving lesson with a concluding whole-class discussion. Nonetheless, the resulting collection of videoed lessons represents a wide range of lesson types, levels of cognitive demand, mathematics content, and classroom discourse patterns.

The IQA, the video analysis instrument used by the MIST project team, was developed to measure aspects of classroom instruction that have been shown in prior empirical study to be related to students' learning in mathematics. The IQA includes ten rubrics organized around a Launch-Explore-Discuss lesson structure, which parallels the organization of the CMP2 curriculum (though they can and often were used on the MIST project to analyze other lesson

structures). That is, lessons and lesson plans in the primary instructional materials used by teachers in districts B and D organized in terms of launch, explore, and discussion phases. The IQA coding schemes assign numerical codes to the level of rigor associated with the mathematics task used (*Academic Rigor of the Task* rubric), the way teachers introduce or launch the task (*Launch* rubric), the ways students are supported to engage in the task (*Academic Rigor of Task Implementation* rubric), and the nature of the concluding discussion which follows student work time (*Academic Rigor of the Discussion* rubric). Though the IQA predates the CCSS-M, the language of many of the IQA rubrics parallel the standards for mathematical practice. For example, the “student providing” rubric is used to assess the nature of student contributions or answers to questions, and whether or not they are justifying their ideas or reasoning conceptually.

Videos were coded by members of the project team in the summers that followed data collection each year. Coders were required to achieve 80 percent agreement with previously consensus-coded videos during the training reliability phase. Additionally, inter-rater agreement was monitored in an ongoing basis, resulting in 15% of the video sample double coded; coders met to agree on any coding discrepancies. Across the three *Academic Rigor* rubrics and across the four years of data collection, percent agreement averaged 70.5%, with kappa scores averaging 0.50. These reliability statistics are comparable to those from other classroom observation instruments used in the MET Project.

**Measure of teachers’ mathematical knowledge for teaching.** Each year, participating teachers took the *Learning Mathematics for Teaching* assessment (LMT). The LMT was designed by Hill et al (2008) to assess teachers’ mathematical knowledge for teaching – i.e., their understanding of mathematics central to supporting student learning. The LMT is a paper-and-

pencil, multiple choice assessment in which teachers are asked to solve mathematics problems *and* assess the validity and pedagogical merit of nonconventional strategies for solving problems. Teachers who were in the study for multiple years were given alternating forms of the assessment that had been found in the past to have equivalent means and standard deviations. LMTs were electronically scored each year and these scores were normalized against the mean and standard deviation of a nationally representative sample collected by Hill et al. Though the confidence intervals on the LMT are too large to compare the depth of one teacher's mathematical knowledge for teaching to that of another teacher, these confidence intervals shrink when comparing a teacher (or a teacher's multi-year average score) against a larger sample (e.g., against the teachers in their district or the larger MIST sample over time).

**Interviews and transcripts.** Each year of the study, participating teachers, instructional coaches, school leaders, and district leaders in participating schools were interviewed by a member of the research team. Interviews were semi-structured, and questions for principals, coaches, and teachers were organized according to the types of support that their districts and schools were implementing (e.g., professional development, instructional coaching, collaborative time, curricular resources), the ways participants were held accountable for certain practices and to whom, and the nature of their work with one another (e.g., teacher-coach relationships, principal-coach relationships, principal-district-leader relationships).

Annual interview training of the interview team (which included between 8 and 15 researchers each year) included instructions on how to ask particular questions, how to probe on participants' responses for adequate detail, and how to interpret and follow up with certain

responses (e.g., district-specific acronyms or programs). Each interview lasted between 45 minutes and an hour, and was audio-recorded and later transcribed.

VHQMI and VSMC, described below, are constructs that were measured through coding of interview transcripts each year. Interview training each year included special attention to these two question sets in an effort to achieve reasonable consistency across interviewers.

**VHQMI.** Interviews with teachers, instructional coaches, principals, and some district leaders (those with more regular or recent contact with math classrooms) included the question, “If you were to observe a math lesson over one or more days, what would you look for to know that the instruction was of high quality?” In addition to this question, the interview protocols used by the research team also included a series of probes specifically designed to elicit the participant’s vision of high-quality math tasks, the role of the teacher in supporting student learning, structures of discourse that support student learning, and how student and teacher questions contribute to high quality classroom discourse. These questions and the rubrics used to code participants’ responses were developed by Munter (2014) to measure participants’ views of what counts as high quality instruction, including the goals for students’ learning, the nature of high quality tasks, the role of the teacher, and the nature and role of classroom discourse, among other aspects of math instruction.

Once interviews were transcribed, a coding team consisting of 2-5 members was trained by the VHQMI rubrics’ developer and coded all transcripts using the NVIVO (versions 8-10) software package. Reliability between coders was maintained at 80% or above for each coding rubric, and 20-25% of interviews were double-coded (see Munter, 2014 for more details on the methods for ensuring reliability across coders and years). Coding of each interview transcript

resulted in scores for multiple dimensions (e.g., *tasks*, *role of the teacher*, and *classroom discourse*, among others) for each participant. Scores ranged from 0 to 4, where lower scores represent a more traditional vision of instruction in which the teacher is the primary source of mathematical knowledge and demonstrator of procedures that students replicate. Higher scores represent a more inquiry-oriented instructional vision in which the teacher serves to facilitate student inquiry using rigorous tasks and conceptual mathematical discourse.

**VSMC.** Teachers' *Views of Students' Mathematical Capabilities* were also measured through interviews. To this end, teachers were asked two primary questions, which were followed by a number of probes: The first question, "When students do not learn as expected, what you find are typically the reasons?" was used to capture teachers' explanations of student struggles or success. The second, "What do you typically do to address these challenges?" was used to capture the nature of supports that teachers either use or would like to use with students who are currently struggling. Ensuing probes asked teachers to specify explanations or types of support for specific groups of students (e.g., ELL's, African American students, or other groups identified by the teacher).

Teachers' responses to these questions were coded by 2-5 members of the research team, in order to assess how teachers framed the sources and appropriate responses to student struggles in mathematics. A teacher's *diagnostic framing* of student difficulties in mathematics refers to their explanations for why students struggle. Teachers' explanations were either coded as productive, meaning they explained student learning or difficulties in terms of the learning opportunities they'd been provided through instruction, or unproductive, meaning they explained differences in student learning in terms of inherent traits of the student or differences in home

support due to parents, income, or community. Teachers who used both productive and unproductive diagnostic framing were coded as “mixed.”

A teacher’s *prognostic framing* of student difficulties in mathematics refers to their views of appropriate supports for students experiencing difficulties. The supports teachers described providing (or wanting to provide) were coded as *productive* when those supports were likely to get students back “in the game” of cognitively demanding mathematics – e.g., secondary math classes whose lessons were designed to anticipate and meet students’ needs for basic skills in upcoming investigations in their primary math class. Supports teachers described were coded as *unproductive* if they were aimed at reducing the cognitive demand of primary math instruction so that they could participate – e.g., providing struggling students with explicit steps to follow in an otherwise novel problem-solving scenario.

**Identification with an ambitious vision.** Teacher interview transcripts were also used to assess the extent to which teachers came to personally identify with the goals of the instructional reforms happening in their institutional settings and in the various support settings they participated in. The primary source of this data came from transcripts of teachers’ responses to an interview question and a number of follow-up probes. Teachers were asked “*In your view, is CMP2 appropriate for all students?*” Teachers’ responses to this question often included their own broad views about the kind of teaching advocated by CMP2, the types of students or classrooms in which this kind of teaching is successful or appropriate, and their opinions about the district’s instructional improvement goals. Teachers’ responses to this question were coded to assess the degree to which they had begun to accept and identify with the vision of math instruction central to the support activities they were engaging with related to CMP2. In addition

to these questions, any other instances in which teachers positioned themselves in alignment with or in resistance to their districts' improvement goals were also used to assess their developing identification with (or resistance to) their district's vision of instruction (explained further in the Analysis Methods section, beginning on page 78).

**Analyzing aspects of teachers' school settings.** In order to investigate the nature of pull-out professional development, instructional leadership, instructional coaching, and TCT that may have influenced the development or decline of the case teachers' instructional practices, I drew on interview data from teachers, coaches, and principals at case teachers' schools. In the following paragraphs, I briefly describe each of these data sources or analytical sources and the aspects of teachers' school settings they were used to investigate.

***Interview transcripts.*** As described above, participating teachers, coaches, and school leaders in the MIST project were interviewed each spring about the nature of school and district supports and accountability for improvement of their practices. Relevant to the present study, teachers were asked each year about:

- the nature of pull-out PD they attended over the course of the year – e.g., what the sessions were about, what activities they did in the sessions, who led them, and whether they found them helpful
- the nature of their one-on-one work with any instructional coaches – e.g., what support the coach provided to them, how often, and whether they found it helpful
- the types of meetings they regularly attended with other math teachers in the school, the activities they did in these meetings, and who led or facilitated them

- their perceptions of their school leaders' expectations for their instruction, and what school leaders looked for or remarked on when they conducted classroom observations

Coaches and school leaders were asked for their perspectives on these same topics, as well as questions about their work with one another. For example, we asked coaches what they were accountable for to their principal, whether they observed instruction with their principal, and what the principal's expectations were for math TCT meetings.

Transcripts of these interviews were used to assess the nature of district pull-out PD and of school-based supports with which each case teachers engaged during the school year. For example, case teachers' descriptions of the activities they typically did with their instructional coaches in one-on-one work were analyzed in order to assess the likelihood that those activities may have influenced changes in the quality of the teacher's instruction over time. My assessments of the quality of these supports were based on my analytical framework, described in the next section.

## **Case Selection**

As described in the previous section on research design, cases for this study consisted of three groups of teachers, whose experiences and perspectives regarding PD, instructional coaching, instructional leadership, and TCT in their schools were the primary data sources for qualitative analysis. Those three groups of cases (referred to throughout as "improvers," "decliners," and "stable teachers") were chosen for study due to their contrasting trajectories in instructional quality.

For the selection of cases for analysis, I worked with Brooks Rosenquist who constructed quantitative models to operationalize my theoretical case selection logic. Together, we first

assembled a dataset that contained all teachers from districts B and D who participated in the MIST project, their annual IQA scores generated through prior researcher coding, and records of which years they participated in the study, and in which school(s).

As noted previously, the IQA coding conducted by the MIST project team contains 10 distinct rubrics which count and categorize different forms of instructional practice and features of the classroom environment, including:

- the cognitive demand of the task
- the degree to which the teacher's launch of the task supports the class to develop a shared understanding of the key contextual features and mathematical relationships in the task
- the level of cognitive demand at which the task is implemented over the course of the lesson
- the cognitive demand of any concluding whole-class discussion(s), which focuses on the degree to which the discussion focuses on conceptual understanding and explanations rather than strictly procedural steps students followed in problem-solving
- the nature of student and teacher talk in the concluding whole-class discussions
- student participation in various sections of the lesson

In order to create useful models of teachers' longitudinal trajectories of instructional quality over time, we needed to reduce this large amount of coding data into a single variable. To accomplish this, we used a composite IQA score which averaged the averages of the three main rubric categories: the quality of the task, the cognitive demand of the launch and implementation

of the lesson, and the quality of the concluding whole-class discussion. Descriptive statistics for this composite score can be found in Table 2 below.

Table 2.	
<i>Descriptive statistics for IQA data in districts B and D, years 1-7</i>	
No. teacher participants in the study for 3+ consecutive years in districts B and D	90
4+ years	37
5+ years	26
<b>Descriptive statistics for IQA scores</b>	
IQA (composite score for best day)	
Mean, SD (whole sample)	2.36, .57
Mean, SD (B and D only)	2.39, .55
Mean, SD (B and D, 3+ years only)	2.41, .55

Using teachers' IQA scores as a primary data source for case selection relies on an assumption that the instruction captured in the MIST classroom video dataset represents the best instruction that participating teachers were capable of providing on an average day of instruction. This assumption brings with it some limitations inherent to measurement error and sampling bias, which I discuss below. For case selection purposes, however, this assumption suggested the use of the "best" or highest-scoring record of instruction available from each teacher. Since the MIST project collected two consecutive days of instruction from each teacher, we thus included only the highest composite score from each teacher each year in the dataset used for case selection.

Next, we used the STATA 10 statistical software package to develop an iterating series of regression models to estimate an average change over all available years of IQA data for each teacher in the restricted sample. This resulted in a line-of-best-fit per teacher, where the slope represented each teacher's estimated longitudinal rate of change (growth or decline) in composite IQA score for available years' data. In other words, a positive slope indicated that the teacher's

instructional quality improved over time, whereas a negative slope indicated a decrease in a teacher's instructional quality declined over time.

Given the significant changes in district policy and supports for teacher learning that occurred during years 5 and 6 of the study, I conjectured that teachers' rates of change might shift direction or steepness during this time period. For example, consider a teacher who, in the first four years of the study demonstrated rapid improvement in her instructional quality. That same teacher might have struggled in the wake of the new standards and curriculum frameworks in district D to locate high-quality tasks that addressed those standards, and instead may have begun using online lesson plans that emphasized fact drills and procedural practice – thus significantly reducing the likelihood that she would engage her students in conceptually-oriented discourse. For this teacher, an average of her composite IQA scores over years 1-7 might be close to 0, but only because her change over years 1-5 and years 5-7 were equally steep but in opposite directions.

To allow for this 'pivot' in year 5, we developed two additional models that paralleled the first, but focused on years 1-5, and then separately on years 5-7. The resulting datasets for all three models were used to select cases, with preference given to the slopes with the greatest absolute value.

Using only teachers for whom we had three or more *consecutive* years of data, I selected the two teachers in each district with the highest and lowest estimated slopes. In the case that there was a tie, I broke the tie based on the number of years the teacher was in the study, opting to include cases with more longitudinal data. The resulting case teachers showed the sharpest rates of improvement or decline in instructional quality over 3 or more years of the study in Districts B and D. See table 3 for a summary of cases.

Table 3					
<i>Summary of Cases Selected for Analysis</i>					
Case type	School	Pseudonym	3 year average change in IQA composite score	Focal years	Years teacher participated in study
Improver	Cypress	Diane	1.590525	2-4	1-7
Improver	Cypress	Cathy	0.6043997	1-4	1-7
Improver	Two Lagoons	Mario	0.795262	5-7	4-8
Improver	Park Falls	Anna	0.636211	2-4	2-6
Decliner	Green Springs	Penny	-1.59053	1-3	1-3
Decliner	Salt Marsh	Li Chu	-1.35195	5-7	5-7
Decliner	Poplar	Karyn	-1.59053	5-7	5-8
Decliner	Cypress	Alishia	-1.03384	5-7	1-7
Stable	Cypress	Roger	n/a	1-2	1-2
Stable	Park Falls	Roxanne	n/a	2-3	2-3
Stable	Two Lagoons	Lois	n/a	5-7	1-7

**Selection of stable procedurally-oriented teachers.** "Stable" teachers were selected for analysis using the following process. I first generated a list of all full MIST participants who taught in the 4 growers' schools (3 schools, since Diane and Cathy were in the same school during the same years) for at least 2 of the years during which the growers' instruction was improving. Next, I examined their IQA scores for those years, and looked for teachers whose instruction was characterized by procedural instruction - e.g., IQA scores of 1s or 2s for task potential, implementation, and discussion. Lessons scored at levels 1 and 2 on the IQA rubrics are generally organized around tasks which focused on procedures without connections or simple fact recall, and classroom discussions (if used) privilege students sharing answers and the steps used to arrive at those answers without reference to the larger mathematical ideas or the students' conceptual understanding of those steps.

It is worth noting that I did not limit the sample of stayers to teachers who were in the growers' school during *all three years* of their improvement. Limiting the sample to teachers with 3+ years in the school during the growers' improvement over-restricted the sample. At Cypress,

the only teachers who were in the MIST sample for those same years of Diane's and Kathy's improvement *were also improving*, though slightly less pronouncedly so. Only one teacher (Roger) taught in procedural ways during those years, and only for two years before switching to become a full-time social studies teacher.

At Park Falls (Anna's school), a pattern of significant turnover between years (both in our sample and in the school) made it difficult to find *any teachers* who were at the school for 3 or more years; in fact there was only 1 other teacher at B6 who was a full participant for the same 3 years that Anna was improving, and her instruction was also improving over those years. There were three teachers who were full participants at Anna's school for two of the three years; two of these either developed or maintained more ambitious instructional practices over the same time period that Anna did, though with less pronounced improvement. The third teacher (Roxanne) maintained procedurally-oriented instruction over those two years, and therefore was selected for analysis.

At Two Lagoons (Mario's school) there were a number of teachers (4) who remained in the sample over the same years that Mario's instruction was improving (y5-7). Of these, three either developed or maintained quite ambitious forms of practice over those years - in other words, either paralleled (though less drastically) Mario's improvement, or had already developed quite accomplished forms of practice and maintained that instructional quality over time. The fourth teacher (Robin) remained procedurally-oriented over all three years that Mario was improving, and was therefore selected for analysis.

While it is encouraging to note that there were few cases of stable procedures-oriented instruction to choose from in the growers' schools, it is not immediately evident that this was due to an overall trend of improvement among teachers in the school. In the case of school Two

Lagoons, for example, most of the teachers in the MIST sample remained in the sample over years 5-7. While many of them were improving in their instructional quality over this time, there were also 4 full-time math teachers in the school who were not included in the MIST sample. Similarly, in Cypress, the MIST participants remained fairly consistent over time; however, there were at least 8 math teachers in the school who were not in the MIST sample. At Park Falls, a large degree of turnover and growth in the school accounted for little consistency in the sample from year to year; in year 2, all math teachers in the building participated in MIST. However, in year 3, there were 2 other math teachers in the building who did not participate, and in year 4 there were 6 non-participating math teachers. Thus, while it could be that these schools did in fact foster a school-wide trend of instructional improvement in mathematics instruction, it could also be that procedures-oriented teachers opted out of participating in the MIST study.

**Potential measurement error and hidden variables in case selection and analysis.**

There is unavoidable potential for error when attempting to measure constructs like instructional quality. Especially when using data to select individual case studies for in-depth analysis, drawing from a limited number of data points exposes the researcher to the potential for hidden variable bias in individual data points. For example, the timing of unforeseeable or uncontrollable events such as snow days or district benchmarking days might disrupt typical classroom instruction on days of video data collection; on the other hand, if data collection occurred on a day immediately following a PD session, the recorded VHQMI or IQA scores might overestimate what is typical for that teacher. Thus each individual data point for MIST participants is subject to some degree of error.

The danger that error of this nature poses to case selection the present study is that teachers' average change over time might not represent actual improvement or decline in their practice, since a change from one year to the next might be explained by the presence or lack of some hidden variable. One way that I addressed this threat was to select teachers who showed improvement or decline between each of the three consecutive years for which they were selected – e.g., the trend was pronounced between each year, not just in the average.

### **Analysis: Framework and Methods**

**Conceptual Framework.** This analysis involved investigating the nature and quality of the various supports for supporting teachers' development of their instructional practice, and the nature of instructional leaders' expectations. While the data analyzed for this study is not sufficiently robust to yield conclusive causal explanations for why teachers' expertise improved or declined over time, the findings of this analysis constitute testable conjectures for how the interactions among various school-based supports and accountability can contribute to teachers' development of expertise related to ambitious teaching. In order to bolster the robustness of these conjectures, my analysis was bounded by a conceptual framework adapted from Jackson et al (2014) which organizes themes across the literature teacher professional development. These principles provided a means of assessing the quality of the supports for teacher learning described by teachers in my sample. Those themes are:

1. Supports for teacher learning should be sustained over time and involve the same group of teachers working together

2. Supports for teachers' learning should focus on issues central to instruction, organized around the materials that teachers use in their classroom
3. Supports should include co-participation with accomplished others in activities that approximate the targeted practices
4. Supports should include pedagogies of investigation and enactment – that is, activity structures which support teachers in understanding the logic behind particular practices and their impact on student learning, as well as activities which support teachers to practice particular moves with the support of a more expert other
5. Supports should include facilitators who press on teachers' ideas differentially and build on their contributions  
  
(adapted from Jackson et al, 2014, pp. 4-5)
6. Supports should focus on supporting teachers to learn high-leverage practices – that is, research-based practices that have significant payoff in supporting student learning, that novices can learn relatively quickly given appropriate support, and that they can incorporate into daily instruction (Grossman, Hammerness, & McDonald, 2009; McDonald et al., 2013).

These principles reflect the current stance from the fields of teacher education and professional development about the nature of supports that are likely to help teachers re-organize their current practices around ambitious goals for student learning. Table 4, below, summarizes how I operationalized each of these principles in my analysis. The last column in Table 4 describes how I operationalized each principle to assess the role of the principal in contributing

to teachers' opportunities to learn by providing resources and supports, and by leveraging aligned instructional expectations.

Table 4.				
<i>Principles of High-Quality Supports for Teacher Learning</i> , adapted from Jackson et al, 2014.				
Principle	Operationalized in pull-out PD	Operationalized in coaching interactions	Operationalized in TCT	Operationalized in role of school leaders
Support is sustained over time and includes the same group of teachers	PD sessions focus on the same practices or topic over multiple sessions in order to give teachers multiple opportunities to learn, enact, and reflect. The teacher attends sessions with the coach or other teachers from his/her school or TCT group, so that there is more opportunity for a continued focus across contexts	The teacher’s work with the [same] coach involves ongoing work focused on developing the same or related practices	TCT meetings have some degree of sustained focus over time (e.g., a focus on incorporating high-level questions in lessons throughout the year/semester, or a focus on improving the quality of launches).	School leaders recognize that teachers sustained support to develop their practice. Their expectations for teachers’ development are consistent across years and reflect a sustained attention to supporting teachers to improve their instruction
Supports are close to practice: organized around critical problems of practice <b>and</b> materials used for daily instruction	Sessions focus on high-leverage aspects of teaching practice and support teachers to use high-quality instructional materials available to teachers on a daily basis	The teacher’s work with the math coach is focused on developing high-leverage instructional practices (e.g., launching tasks without reducing the cognitive demand)	TCT activities include a focus on learning or improving high-leverage instructional practices, including making use of instructional materials like CMP or assessment data in ways that promote ambitious instruction	The principal’s expectations for teachers’ instruction are aligned with an ambitious vision of instruction, and their goals for teachers’ work in TCT and PD include focusing on the development of high-leverage practices.
Supports include co-participation with more expert others	Sessions are led by PD leaders who are accomplished in ambitious instructional practice	The math coach is relatively accomplished in knowledge and practice mathematics instruction	The teacher’s TCT group is led by (or at least includes) someone who is relatively accomplished in ambitious instructional practice	The principal hires or promotes a coach with instructional expertise (or asks others for help in doing so). When assigning the coach duties, the principal provides prioritizes their work in supporting teachers one-on-one and in group contexts (rather than assigning them administrative or tutoring duties)

<p>Support includes pedagogies of investigation and enactment</p>	<p>Sessions include pedagogies of investigation and enactment, such as doing math together, rehearsing practices together, and reflecting on their own practice</p>	<p>The math coach supports the teacher to reflect on their own practices, investigate practice by observing the coach teach students, or enact practices with support – e.g., through co-teaching</p>	<p>TCT activities include pedagogies of investigation and enactment, such as doing math together, rehearsing practices together, and reflecting on their own practice</p>	
<p>Supports are focused on high leverage practices</p>	<p>Across these supports and school leaders' expectations, the overarching goal is to support teachers to improve the quality of classroom learning opportunities for students. There is consistency across these various supports in the practices that teachers are to develop, and whether or not they are held accountable for implementing them.</p>			

**Analysis Methods.** My analysis process was motivated by the goal to understand and explain each case teacher’s improvement or decline in instructional quality, test these conjectures by comparing them with other cases, and identify contingencies in these conjectures by examining potentially disconfirming cases of stable teachers who did not improve though they had the same supports available to them as improvers. My general approach to coding for this analysis was to first use the available data and prior analyses from each teacher to familiarize myself with the teacher’s experience as a math teacher in their school, including but not limited to their access to supports and their perception of their school leaders’ expectations for their instruction. I used an iterative process of open coding and unstructured memoing to collect emergent themes and new research questions that came out of the data. I then moved to more structured coding and memoing that would allow me to make comparisons within and between teachers and groups of teachers. All coding was done using NVIVO 10 Qualitative Coding software. Below, I detail the phases of my coding process, referring to the relevant codebooks, templates, and data tables used for my analysis process (Appendices A-E).

***Phase 1: Open coding and memoing.*** For my first reading of each teacher’s interview transcripts, I used NVIVO 10 to conduct paragraph-by-paragraph coding at the descriptive level that would provide a visual summary of the topics and ideas represented in the interview. When a paragraph addressed multiple topics or major ideas, I used multiple codes (e.g., “CMP was forced on us,” “CMP not good for low readers.”) Following this open coding, I wrote an “Open-ended Themes” memo for each transcript, and then for each teacher’s transcripts across years. These memos were focused loosely on the supports and accountability that were the focus of my research questions, but also included additional themes and questions that emerged out of the data – e.g., “*Cathy considered herself an ‘outsider’ in her first year, but came to feel part of the*

8<sup>th</sup> grade team by her 4<sup>th</sup> year at DI....” These memos were not structured using templates, but instead were left open to unanticipated themes that might emerge within or across cases.

**Phase 2: *A priori coding and summarizing.*** Following this phase of emergent coding and memoing, I then used a set of *a-priori* codes to categorize the form and nature of supports and instructional expectations each teacher described in interviews. I used these same *a-priori* code to code the school summary forms from that teacher’s school/year, in order to triangulate (when possible) the teacher’s reports with those of other teachers in the school about the nature of TCT activities, the coach’s work, and the principal’s expectations for teachers’ practices. The codebook for these *a-priori* codes appears in Appendix A.

Following this process of *a-priori* coding, for all years’ interviews for each teacher, I summarized the teachers’ reports (and or changes in those reports) over time in a Case Summary Form (see appendix B). These memos were both descriptive and interpretive; I drew on my coding for each teacher’s transcripts, as well as my conceptual framework, to make assessments of the quality of each support and/or the nature of school leaders’ instructional expectations, and the likelihood that these supports/expectations may have contributed to the teacher’s instructional improvement, decline, or stability over time.

**Phase 3: *Constant comparative coding of identification/buy-in excerpts.*** While coding the transcripts of improvers and decliners using my *a-priori* codes, I also grouped data that would be relevant for analysis of their identification over time. Any instances in which the teacher explicitly positioned themselves vis-à-vis CMP2, an ambitious vision of instruction, the district’s reform goals, or their principal’s improvement goals, I coded the entire turn of talk using a single code: “*Buy-in or identification excerpts.*” Once all improvers and decliners had been coded using my *a priori* codes, I returned to the excerpts coded to ‘buy in or identification

excerpts' in order to group similar turns of talk together. In particular, I wanted to understand themes across teachers in why they considered CMP appropriate or inappropriate for their students. Using the constant comparative method (Glaser & Strauss, 1967), I grouped teachers' responses by theme and similarity, then re-read back through these groups to ensure analytical 'fit' between excerpts, and to make sure these excerpts seemed fairly representative of the teachers' larger stance toward CMP2.

Grouping these codes resulted in four primary categories of reasons for why teachers viewed CMP as an appropriate curriculum for all students (or not). There was only one teacher who never volunteered an opinion about CMP over three years of interviews (Anna, an improver). Of the remaining 6 improvers and decliners, all teachers identified particular aspects of CMP that make it difficult to teach, or challenging for particular students. Some teachers described these as challenges that could be mitigated with a measure of hard work; others did not.

The major themes in teachers' responses were:

1. CMP is or is not appropriate because of the way it presents mathematics
2. CMP is or is not appropriate because of the kinds of challenges it presents to students
3. CMP is not appropriate because it does not align to the state standards and/or state tests
4. CMP requires a lot of work as a teacher, which [is/is not] worth the effort.

Overall, there were no pronounced differences between groups (improvers/decliners) in whether/how they appealed to these reasons for CMP's [in]appropriateness. E.g., improvers were no more likely than decliners to explicitly discuss problems/affordances of CMP in terms of how the curriculum presents/sequences mathematics content.

***Phase 4: Synthesis and memoing of case teachers' buy-in/identification.*** After identifying these themes, I then returned to each individual teacher's group of transcripts in order to identify trends or changes over time in how that teacher positioned him/herself in relation to the district's reform efforts, an inquiry-oriented vision of instruction, and CMP in particular. I looked for themes that appeared most frequently across the teacher's transcripts. For example, across years, Penny repeatedly referred to theme 2, that CMP2 was not appropriate because it did not closely align to the state standards. Cathy, on the other hand, maintained a positive view of CMP2, but identified new challenges over years 1-3 that she had to work through in order to make it work with her students.

These themes, and factors that were particular to the teacher's school (e.g., a principal who was adamant/lax about the use of CMP2 or particular instructional practices) were then summarized in a "Buy-in/Identification Memo" (see appendix C). Thus, by the end of phase 4, I had accumulated 11 open-ended themes memos, 11 case summary memos, and 11 buy-in/identification memos, representing the four improvers, four decliners, and three stable teachers analyzed for this study.

***Phase 5: Comparison of Improvers with Decliners.*** Using NVIVO 10, I produced a number of data tables comparing the a-priori coding of improvers with decliners – pooled over time, and organized by year. These tables were useful for getting a top-level view of differences between the school settings of improvers and decliners, as were comparative tables which organized prior assessments of teachers' instructional expertise and changes therein (see Appendix D).

***Phase 6: Comparison of Improvers with Stable Teachers.*** Drawing on my case summary forms, open-ended themes memos, and buy-in/identification memos, I looked at each stable

procedures-oriented teacher's experience with supports/instructional expectations in comparison with those from the improver in their school. I also compared the extents to which each teacher in the pair came to identify with the district's goals for implementing CMP2 and inquiry math instruction more broadly across middle schools. I wrote a summary of these comparisons in a memo detailing similarities and differences in each pair in an "Improver vs. Stable Teacher" memo (see Appendix D).

***Phase 7: Triangulation and Validity.*** Given the forms of data collected by the MIST study, I was able to draw on additional data besides teachers' interviews in order to triangulate their reports about school-based supports and professional development opportunities (including pull-out PD, coaching, and TCT). For example, teachers filled out a survey each spring in which they described how frequently they engaged in particular coaching activities (such as modeling, co-teaching, etc.), the focus of the PD they had attended during the year, and how frequently they met with other teachers in their school to collaborate. For their reports about their activities in TCT, it was possible to triangulate teachers' descriptions with other participants at the same school (using the school summary forms written by other MIST researchers). To some extent, it was also possible to triangulate (or at least compare) teachers' perceptions of their school leaders' expectations with other teachers in the school, and with the principal's own reports of their expectations (though the latter was often different from what teachers perceived).

For the more interpretive codes used to categorize teachers' developing identification/resistance vis-à-vis the district's reform efforts, I had only limited additional data to draw on: teachers' response to a survey item in which they answered how consistent CMP was with their own beliefs, and (occasionally) other teachers' reports of the views of others in their schools. However, the former offers very little insight into the reasoning behind their response

(e.g., is CMP inconsistent with their mathematical beliefs? Their pedagogical beliefs? Their beliefs about teacher autonomy?), and the latter depends greatly on chance. For example, at Cypress, there were only two 6<sup>th</sup> grade math teachers in MIST year 2: Diane and Roger. Therefore, when Diane refers to “the other 6<sup>th</sup> grade teacher” and his distaste for CMP2, it can be safely assumed that she is describing Roger. However, these chance references were rare.

Thus, to introduce means of interpretive triangulation, and measures of reliability into my analysis, I followed the suggestions of Yin (2003) and Stake (1995) to first establish a clear “chain of evidence” and then engage an outside researcher in both *descriptive triangulation* and *theory triangulation* (Stake, 1995, p. 113). Descriptive triangulation involves determining whether an outsider to the study would observe similar phenomena when posed with the same question; theory triangulation involves determining whether an outsider would interpret those phenomena similarly when working from the same conceptual framework. This involved recruiting an additional member of the MIST research team (Denise Kelly) who was familiar with the case study contexts, my analytical framework, my research questions, and my case selection process, to examine the interview data of three of the 11 case teachers (27% of the cases). The goal in involving an additional coder was to see whether, when attempting to answer the same questions about each teacher’s identification with their district’s reform efforts, she would (1) identify the same interview excerpts as relevant and bearing on the questions, and (2) draw similar conclusions about the teacher from those excerpts.

To facilitate this triangulation process, I first oriented Denise to the goals of my study, to my research questions, and to my case selection process. I then provided her with a procedure for reading through all of the interview data from Diane (4 years of transcripts), Roger (2 years), and Mario (4 years), and provided her with templates for memos which asked her to summarize (and

describe evidence for) her conclusions about these participants' sense of identification with their district's reform efforts (see Appendices F and G). During this orientation, I did not provide Denice with any examples of the evidence/excerpts I had drawn on, nor with any of my own interpretations of that data.

After she finished her analysis of Diane's and Roger's transcripts, we met to discuss her memos and compare them with my own assessments of Diane's and Roger's identification with CMP2 and District D's reform efforts. We then repeated this process with Mario's data. Our separate analyses resulted in nearly identical conclusions about these teachers' identification with CMP2 and their districts' reform goals. The only difference between our analyses was the number of specific excerpts we both identified: Denice identified several additional, confirming excerpts that I had not specifically cited in my own memos.

The findings of this triangulation process do not preclude the possibility that other data or other conceptual frameworks might have yielded different interpretations of these teachers' teaching identification. However, these findings do lend robustness to the use of my conceptual framework when analyzing these particular data.

**Limitations.** The scope of the present study is intentionally narrow in that it is focused on outlying cases of teachers whose instruction improved or declined over time. This limited scope was purposeful: by selecting outlying cases and exploring the interrelationships between individual teachers' expertise, the nature and quality of instructional supports which were aimed at influencing their practices, and the instructional expectations of school leaders, this study will contribute and explore hypotheses related to school settings that support teachers to develop ambitious practices. While a small number of cases allows for more thorough analysis of the

available data, these case studies are necessarily limited in their generalizability. Given the sample from which this analysis draws and the larger goals of the MIST project, the district settings in which these teachers work are not representative of all urban school districts at the time data were collected. Given the ambitious goals and extensive supports that these districts had in place for teacher learning, they constitute very unique settings. However, given the rigor of the now widely-implemented CCSS-M, and the challenges inherent to developing ambitious classroom instruction, the findings of this study are generalizable to teachers in other schools considering the potential contribution of school-based supports for teacher learning.

Another limitation to the present study is that this analysis was constrained to the data that were previously collected in the MIST project. This meant that member-checks and additional data collection were not possible. Additionally, it meant I was constrained to the *scope* of the MIST project data and the questions that the research team originally set out to answer. For example, one potential explanation for a teacher's instructional improvement over time could be their concurrent enrollment in a graduate education program focused on math instruction. Though we did ask teachers about their education and coursework in math education on the annual teacher survey, teachers were not asked to describe the kinds of activities they did in these courses or the expertise of their instructors. Thus, unless the teacher volunteered the information in an interview, I did not have any way of knowing whether outside educational opportunities may have contributed to their instructional improvement.

## CHAPTER IV

### FINDINGS

#### Introduction

In this chapter, I present the findings of my analysis, organized around my two overarching research questions:

1. How are school-level learning opportunities and instructional expectations related to the quality of teachers' instruction over time?
2. How do teachers' instructional expertise and personal identification with the goals of their districts' reform efforts mediate the influence of instructional supports and expectations on their development of ambitious practice?

As described in the methods chapter, answering these questions involved a comparative analysis of 11 cases of teachers in the MIST sample whose instructional quality improved, declined, or remained stable and procedurally-oriented over three or more years. In this analysis, I explored conjectures relating teachers' likelihood to develop ambitious instruction with the availability and quality of three primary forms of support (one-on-one coaching, teacher collaborative time, and pull-out PD) as well as their perceptions of their school leaders' instructional expectations. I also explored conjectures relating teachers' development to their current forms of expertise – namely, their instructional quality, their VHQMI, their MKT, their views of students' mathematical capabilities, and the extent to which they came to identify as inquiry-oriented math teachers.

As I will unpack in the following sections, the key findings of my analysis are that (1) ongoing, one-on-one support from an instructional coach with expertise in math instruction can serve as a high-leverage support for teachers to develop ambitious instruction, that (2) regardless of the form or quality of school-based supports for teacher learning, teachers are unlikely to develop or sustain ambitious forms of practice if they do not come to identify with an inquiry-oriented vision of instruction, and that (3) this identification is underpinned by an ambitious vision of instruction and the view that such instruction is appropriate for all of their students.

In the first section below, I present findings from my analysis of the four cases of atypical *improvement* selected from the larger MIST dataset. This section lays out cross-cutting similarities, when they exist, in the forms and quality of the district and school conditions in which these teachers worked during the years of their improvement. When describing these forms of support, I also provide additional or contrasting examples of the school settings of those teachers whose instructional quality *declined* in order to elaborate on my findings about how these supports might relate to teachers' instruction over time.

In the second section, I present findings from my comparative analysis of the expertise and identification of the atypical improving and the stable procedurally-oriented teachers in their school during the same years, focusing on how differences in teachers' initial expertise and identification may account for differential engagement with school-based supports for teacher learning, and ultimately, different forms of instruction. Table 5, below, presents some descriptive information about the eleven improvers (I) decliners (D) and stable procedurally-oriented (S) teachers.

Case type	Pseudonym	Years expc.	MIST years	Dist	School	Gr. taught	Role/participation changes
I	Diane	5	2-4	D	Cypress	6	New to her school in year 2; became math coach in year 5
I	Cathy	6	1-4	D	Cypress	8	New to her school in year 1
I	Mario	2	5-7	B	Two Lagoons	6	Brand new T in year 4
I	Anna	4 (1 in math)	2-4	B	Park Falls	8	Taught history before, became math lead in year 3
D	Penny	30	1-3	B	Meadowbrook	6	Retired at the end of year 3
D	Li	1	5-7	B	Salt Marsh	6	Brand new teacher in year 5
D	Karyn	3	5-7	D	Poplar	6	
D	Alishia	19	5-7	D	Cypress	8	Had been at Cypress for MIST y 1-4 also; Retired in Jan 2015 (mid-year 8)
S	Roxanne	3 (1 in math)	2-3	B	Park Falls	8	Taught science before, first year teaching math in year 2. Left school after year 3.
S	Roger	4	1-2	D	Cypress	6	Left study after year 2; taught social studies
S	Lois	6	5-7	B	Two Lagoons	7	

Note: years of experience refers to how many years the teacher had been teaching at the start of the MIST study years in column 4, when the teacher's atypical improvement or decline began (or over which they were consistently procedurally-oriented). Because interviews were conducted in January, I included the Fall as teaching experience; thus "1 year experience" here (for Anna, Li, Roxanne) indicates that they were in the middle of their first year teaching (or teaching math).

### **Part 1: School-Based Supports and Expectations for Improvers' Development**

Based on my review of the literature on factors which contribute to teachers' instructional improvement, my comparative analysis of improvers and decliners focused on four key forms of support and accountability: teachers' access to and work with instructional coaches, their

participation in pull-out professional development, their opportunities to learn in TCT, and their perceptions of their school leaders' expectations for their instruction. I analyzed case teachers' reports to assess whether and how these supports and expectations may have contributed to their changes in practice over time.

**One-On-One Work with Instructional Coaches.** Throughout years 1-7, most teachers in Districts B and D had some access to an instructional coach who either worked in or visited their buildings. However, the extent of coaches' work with teachers, and their expertise in ambitious mathematics instruction varied greatly depending on the district, the school conditions, and the year.

*District coaching initiatives.* In District B, the district provided funding for each school to have a part-time, school-based coach, who usually taught half the day and were to work as a coach the rest of the day. Once a month, these coaches gathered at the district office for meetings and professional development from the district math specialists. These coaches were selected by their principals, and were ultimately held accountable to their principal's expectations for their work. This meant that the ways school-based coaches in District B worked with teachers varied greatly from school to school; in some schools, coaches served highly administrative roles, managing paperwork and aggregating student assessment data in order to manage programs for struggling students, such as pull-out tutorials for subgroups of failing students. In other schools, coaches worked closely with teachers in cycles of co-planning, observation, and feedback. Across most schools in District B, these school-based coaches also served as a liaison between the district and the teachers, bringing back from their meetings information about upcoming PD, policies, or available instructional resources to teachers in their school.

In year 6, after seeing little measurable impact of this coaching structure, the district (along with several other district-level reorganizations) changed it. In years 6-7, District B provided three district-based coaches who were assigned to work with multiple schools at once. These three coaches had considerable expertise in inquiry-oriented instruction, including relatively advanced MKT, a sophisticated VHQMI, and very productive views of students' math capabilities. In addition to rotating between multiple schools (often prioritizing schools with the lowest student achievement scores/progress), these district math coaches also led several district PD sessions each year. These coaches were accountable to both district leadership directors (e.g., assistant superintendents who oversaw a number of schools in a zone or area) and to district math specialists.

Coaching (and changes to coaching) took an almost opposite form in District D. In years 1-5, the district had several math specialists who each worked with a large number of schools; these coaches were, when compared with the larger MIST sample of teachers, relatively more accomplished in vision and MKT. Similarly to the coaches in years 6-7 in District B, they were stretched thin across many schools, and often prioritized their work in schools with large numbers of new teachers, or in schools with especially low student achievement on state tests. When they visited schools, they often worked one-on-one with teachers in their classrooms and attended TCT to answer teachers' questions about curriculum or instruction. While these coaches usually reported back to the school principal about what they had done during their visit, they were ultimately accountable to district math specialists.

In year 6, a change in senior district leadership and the district's agenda for improvement resulted in the reassignment of all district-level content coaches to become school-based coaches who were accountable to the school principal. Since there were only a limited number of these

coaches in each content area, not every school received a coach with expertise in math: Karyn's school, for instance, received a former district social studies coach. The district advised that these coaches should contribute to the improvement of student achievement across the school (rather than in any one content area), and that they should facilitate the organization of PLCs in the building. Ultimately, though, the details of their work were left up to their principals.

*Findings.* When compared with decliners, three improvers (Diane, Cathy, and Mario) had ongoing access to instructional coaches who were relatively accomplished in inquiry-oriented instruction, or had at least a substantially more sophisticated vision of inquiry-oriented instruction, and who had productive views of students' capabilities. Their work with their coaches included both formal coaching activities, and informal advice-seeking. For example, all three described going to their coaches with questions or for advice about math instruction, and they all said that their coaches frequently observed their instruction, provided feedback, and occasionally modeled instruction. In all three cases, teachers saw their coaches as supportive colleagues who were not there to evaluate or punish them based on their classroom observations.

At Cypress Middle School, where Diane and Cathy worked, the principal began hiring and recruiting new teachers in MIST year 1 with the explicit purpose of attracting teachers who were interested and possibly accomplished with using CMP2, the district's newly-adopted curriculum. In year 2, he used school funding he had been 'saving up' for several years to hire a math coach who would be charged specifically with supporting teachers (including some who were unsure or skeptical of CMP2) to implement the curriculum effectively. Karon, the coach he hired, was a former high school math teacher who admitted that she'd had no experience with the specifics of the CMP curriculum herself, but understood and agreed with its approach of presenting math concepts through investigation and inquiry rather than demonstration and

practice. She quickly developed a very sophisticated vision of inquiry math instruction, and had substantially higher MKT than most of the teachers in the building. Furthermore, she was committed to learning more about CMP2 and helping the teachers at Cypress to become comfortable with the curriculum and use it effectively with their students. Both Diane and Cathy worked frequently with Karon when she started at Cypress – in Diane’s view, because Karon wanted to quickly learn alongside teachers who were “onboard” with the curriculum. Karon co-planned, co-taught, and looked at student work with them, modeled lessons, and frequently observed their instruction. In subsequent years, Karon turned to focus more on teachers who were frustrated or struggling with CMP; even so, she continued to offer support and advice to both Diane and Cathy, spending a week at a time with each math teacher in the school to provide one-on-one support. In years 3 and 4, these week-long rotations also included one or more pre-planned periods in which Karon would serve as a substitute so that teachers could observe other teachers’ instruction.

Unlike Diane and Cathy’s work with their coach, Mario’s coaching support was not from the same coach over all three years that his instruction improved. In his first year as a teacher at Two Lagoons MS (MIST year 4), Mario frequently sought advice about math instruction from his assistant principal, Vera, who had formerly been a district math specialist, and was particularly expert in the CMP materials and how to make changes to the investigations without compromising on the mathematical goals for student learning – something Mario noted was not always true of his colleagues’ adaptations to the curriculum. He frequently went to Vera for advice or support specifically about teaching CMP, and to help him make sense of how to integrate other practices, such as student grouping strategies, into CMP investigations. In year 5, Mario did not receive any one-on-one support from a coach to improve his instruction; that year,

in his own words, he “rebelled” against CMP and pieced together his own curriculum from online resources; his IQA scores from that year indicate that the tasks he chose or constructed were very procedural in nature and had little focus on students’ conceptual understanding. This “rebellion” ended the following year as Mario realized he didn’t have time to create a new curriculum himself every day, and he decided that in the end, CMP2 was the best resource available for his students to develop “processing” or understanding in mathematics beyond just getting the right answers. In years 6 and 7, Mario’s school was assigned a district math coach, Malcolm, who had substantial expertise in math instruction. By year 7, Mario was working with Malcolm regularly to co-plan instruction and get Malcolm’s feedback on his teaching after observations. As Mario worked more closely with Malcolm over years 6 and 7, his instructional quality improved consistently.

The fourth improver, Anna, did not receive any regular one-on-one coaching support that we know about, in part because she *became* the school-based coach at Park Falls MS her second year. As the coach, Anna taught half-days (2 periods instead of 4 or 5). The rest of the day, she was expected (by the district) to work with individual teachers and groups of teachers to support their development and implementation of CMP2, and was also expected (by her principal) to coordinate and plan for a number of extensive tutoring programs that took place before, during, and after the school day. As a coach, Anna also attended monthly meetings at the district office with district math leaders. Anna described these sessions as a mix of information-sharing (e.g., distributing district math data to take back to their schools) and coaching development (e.g., doing a book study on effective coaching methods that they could use with teachers). While she did say these meeting (the book studies, in particular) were helpful for her work as a coach, her descriptions of the central activities of these sessions are limited. Interviews with the district

math leaders who organized and facilitated the sessions indicate that they involved three or more half-day discussions of books about cognitive coaching more generally, and math coaching in particular, and also included role-playing opportunities in which coaches took turns rehearsing particular coaching strategies with other coaches acting as teachers. Though Anna's descriptions do not provide much detail about how she engaged with district math leaders at these monthly meetings, we do know that the math leaders who facilitated these discussions had significant expertise in mathematics instruction and in teacher learning – such as Keisha, whose VHQMI particularly sophisticated and who had worked for several years as a math coach before taking the district position. It is possible that through these monthly sessions, along with informally seeking advice from Keisha (as she described doing regularly in years 3-4), Anna came to identify with an ambitious vision of instruction, or at least developed a sense of obligation to the district to implement CMP2 as Keisha and others intended.

In summary, all four improvers had access to school or district coaches with significant expertise in math instruction, and three of the four specifically described getting one-on-one support from these coaches as they implemented CMP2 in their own classrooms. In the latter three cases, the teachers' work with their coaches was focused on improving their own practice in implementing CMP2, and included high-leverage coaching activities such as co-teaching, modeling, and observations and feedback.

By contrast, the decliners either lost or, in one case, never received support from a math coach with relatively more instructional expertise. In the early years of their decline, when their instructional quality was still quite high, three of the four decliners (Karyn, Penny, and Alishia) described getting support from a math coach with significant expertise in CMP, productive views of students' mathematical capabilities, and a sophisticated vision of inquiry-oriented instruction.

All three described engaging in some potentially high-leverage coaching activities with their coach (e.g., being observed and receiving feedback; watching their coach model a lesson; attending school-based math PD led by the coach). However, over the years of their declining instructional quality, either the coach was replaced with someone with less expertise, the activities they did with their coach became low-level, or both. The fourth teacher, Li, was never supported by a school-based math coach, and only received support from an expert district coach for one year during his three years of decline (the second year).

In Karyn's case, her school had been supported by a district math coach who visited the school a couple days a week to attend math department meetings and grade-level meetings, and provide one-on-one coaching. In year 6, the district leaders in District D decided to reassign all district-based content coaches to become school-based, generalist coaches whose work would be focused on supporting the principal to improve achievement scores, and facilitate the organization of PLCs in the school. In Karyn's case, the coach assigned to her school had been a district social studies specialist. She had very low MKT, which was not surprising, though her VHQMI was on-par with the teachers in the school. Given her lack of expertise in math content, she admitted to having no insight into matters of content or curriculum, and didn't work one-on-one with any teachers in the math department.

In Penny's case, the math coach they had initially was relatively accomplished in inquiry but left after year 1 to become a district math specialist. She was replaced by a coach who Penny had very little respect for personally or mathematically, and whose instructional expertise was no more sophisticated than the other math teachers in the building.

In Alishia's case, the decline in coaching support was not due to the coach being reassigned, but changes in the kinds of coaching activities she was doing. Her coach (Diane, an

improver who became a coach at Cypress in year 5), was quite accomplished in ambitious instruction, and at first prioritized her work with teachers as a support for their learning. Subsequently, Diane was given an increasing number of administrative, data-related responsibilities in years 5-7, and spent less and less time actually working with teachers (including Alishia) except to prepare/provide materials for remedial tutoring periods for ‘bubble’ kids.

Li, who didn’t have an expert math coach to begin with, did get some focused one-on-one support from a district math coach in his second year teaching, but this support disappeared in his third year as the coach began to focus more on less experienced teachers in the building.

**Pull-out professional development.** Due to the retrospective nature of interview data, case teachers’ interviews were not optimal for assessing the quality of PD that teachers attended over the prior year (e.g., the nature of specific activities teachers did during sessions; the expertise and facilitation skill of the leaders), and the scope of the larger MIST study did not allow for extensive data collection around PD that took place in each district over years 1-7. However, case teachers’ interviews were helpful for assessing the amount of PD they attended, the degree to which it was focused on mathematics in general or CMP units/investigations in particular, and whether or not the teachers found it helpful. While teachers’ self-reports may not be a robust measure of their learning (Dunlap, Hunter, Rosenquist, and Smith, revise and resubmit), their descriptions of the sessions they attended and the degree to which they found them helpful did provide insight into whether and why teachers viewed certain sessions as useful, and whether they were motivated to attend subsequent sessions.

***District PD initiatives.*** The PD that teachers described included extended (often multi-day) sessions that took place over the summer, and pull-out or after-school sessions that occurred during the school year. In both districts, the most time-intensive math-specific PD sessions teachers described were week-long CMP “institutes” that were held during the summers prior to MIST years 1-3. These institutes were either held at the Michigan State University campus and led by teachers or coaches with significant experience and expertise using CMP with students, or were hosted locally and led by state or district math specialists. In District D, a group of teachers who had been identified as early adopters in a “cohort” model in MIST years 1 and 2 were involved in these district-based CMP institutes in years 2 and 3, and then facilitated half-day sessions at their schools specific to using CMP in their own grade levels.

In the early years of the districts’ adoption of CMP2, these PD institutes usually involved first familiarizing teachers with the instructional materials, their organization, and how to use and sequence the resources. Teachers were then led through a number of investigations themselves, working tasks that they were to use with their students. Based on teachers’ descriptions, these sessions seemed to involve roleplaying as students in both group work and whole-class discussions, with facilitators demonstrating discussion orchestration practices.

District professional development that occurred during the school year varied between districts and years. In general, math-specific PD in both districts tended to focus on the immediately upcoming CMP units or tasks (both districts in years 1-4); specific instructional practices such as *effectively launching tasks* (year 5 in both districts) and *leading whole class discussions* (year 6 in District B); and how changes in state math standards would impact curricula, sequencing, and assessments (particularly in years 5-6).

In both districts, these math-specific sessions were usually interspersed with more content-general sessions that reflected the different foci of two districts’ improvement strategies each year. For example, District B offered sessions on Accountable Talk (y1-4), on strategies for working with ELL students (most years), on flexible grouping strategies (y5-6), and on student diversity (y2-3). District D also included sessions on working with diverse students (y2-4), sessions about strategies for supporting Special Education students (most years), integrating technology into instruction (years 2-4), and formative assessment lessons and strategies (y 3-7). Both districts also offered a number of sessions during years 1-7 that did not focus specifically on instruction, but rather on topics such as using district tools for creating assessments or viewing student assessment results, or revisions to state policy regarding special education students.

Table 6.			
<i>Summary of teachers’ participation in math-specific PD during their years of improvement or decline.</i>			
Case type	Pseudonym	Reported attending 1 or more summer CMP institute	Reported attending other pull-out PD on ambitious math practices or curricula
Improvers	Diane	Yes	Yes – 6 <sup>th</sup> grade CMP cohort member
	Cathy	Yes	Yes – 8 <sup>th</sup> grade CMP cohort member
	Mario	No	Yes – launching tasks, just-in-time CMP PD
	Anna	Yes	Yes – coach sessions on observing teachers and providing feedback; math coach book study
Decliners	Peggy	Yes	Yes – just-in-time CMP PD
	Li	No	Yes – launching tasks, just-in-time CMP PD
	Alishia	Yes; previous to decline years	Yes – using formative assessment lessons
	Karyn	Yes; previous to decline years	Yes – launching tasks, using formative assessment lessons

**Findings.** Table 6 summarizes teachers’ participation in CMP institutes and/or PD focused on ambitious math practices or curriculum in general. In contrast to the differences

between improving and declining teachers' access to high quality one-on-one coaching, I was unable to detect any major differences in either the content or quality of the PD that improvers, decliners, and stable proceduralists reported attending. Nearly all attended PD each summer, and many teachers in District B also attended PD sessions regularly during the school year ("Stipend" sessions).

Of the 11 case study teachers, only 2 (Mario and Li) did not attend a week-long CMP2 institute, in both cases because by year 5 District B was no longer providing these intensive summer sessions, and Mario was a late-hire in year 4 who missed the session that summer.

Importantly, some of the teachers across groups attended the same institutes. For example: Penny (decliner), Anna (improver), and Roxanne (stable) all attended summer CMP institutes in the summers of 2008 and 2009. All three described these sessions positively; Anna, for example, said the prolonged time spent working through the CMP investigations helped her understand "the importance of going through each investigation, doing it by yourself so you can know what kinds of questions your kids are going to have, and you'll be prepared for those" (Anna, year 2 interview). Peggy described both the institutes and the presenters as "excellent" (Penny, year 3 interview). Diane and Roger (improver and stable procedurally-oriented teacher, respectively) both attended summer CMP institutes provided by the district. Interestingly, though they both found these institutes helpful, their descriptions of what they got out of them were quite different. For Diane, who had been using CMP2 for at least a year, the most beneficial part of the institutes was to have time to consider both student thinking and pedagogy, by working through investigations as a student before planning the lesson as a teacher. She valued the opportunity to "sit through an actual class where [PD leaders] teach you, and you see it done" (Diane, year 2 interview). In particular, she appreciated that the facilitators at these institutes had

been teaching CMP themselves for long enough that they could answer questions and provide practical advice about how to modify the materials when students were struggling. Roger, on the other hand, was more interested in becoming familiar with the actual materials – how they were organized and so forth. “I think in the beginning [the CMP institute] might have influenced how I used it, but more so it was just to help me get acquainted with the book” (Roger, year 1 interview). In contrast to Diane, Roger felt that the facilitator “strung [the teachers] along, took a lot longer than it needed to take” (year 1 interview). Their different perspectives on the CMP institutes and their usefulness provide evidence that these teachers’ goals for their own teaching, and their professional identities as teachers, likely shaped their engagement with PD and other supports that were offered to them throughout the year. I return to this topic in the next section, when I present the findings of my comparative analysis of improvers with their stable, procedural colleagues.

Nearly all the case study teachers described attending one or more PD sessions held during or after the school day each year; some of these were specific to math instruction, while others focus on more general topics, like using graphic organizers or strategies to use with ELL or Special Education students. When compared with teachers’ opinions of the summer institutes, there was far more variation in teachers’ opinions of the usefulness of these sessions. In general, both improvers and decliners found sessions helpful when they were organized around upcoming units or investigations and allowed time to work through and plan for upcoming instruction. In particular, teachers appreciated sessions that included time to prepare for an upcoming lesson or unit, or improve upon their existing plans. Penny, for example, appreciated that in year 1 when the district was transitioning to CMP2 in all middle schools, the district PD sessions provided time to “talk about the upcoming lessons, which is good because we’re not familiar with the

upcoming lessons, and what would you do to improve upon the lesson that's coming" (Penny, year 1 interview).

At the same time, both improvers and decliners also expressed dissatisfaction and even frustration with many of the pull-out sessions that occurred during the school year. The most frequent complaint across groups was that sessions often focused on topics that were not immediately relevant to their work. In some cases, teachers took issue with the goals of the sessions they attended. For example, Penny described a series of sessions she attended in year 3 as "useless" because the sessions focused specifically on the Algebra strand. "We really and truly do not do much with Algebra in the 6<sup>th</sup> grade, and that was the focus...I mean, we, I would say it was not nearly as valuable as other [sessions], and I've heard that from anyone else who attended..." (year 3 interview). For her part, Karyn expressed frustration in year 5 after attending a three-hour session that focused on creating and planning a lesson on particular standards because she discovered the same lesson had been created and made available beforehand on the district's online curriculum guide. In both cases, the leaders of these sessions may have seen potential benefits to spending time on these PD activities, such as preparing 6<sup>th</sup> grade teachers like Penny to integrate pre-algebra skills into their curriculum to prepare students for 7<sup>th</sup> and 8<sup>th</sup> grade, or supporting teachers like Karyn to develop productive planning routines in advance of the district's movement to the Common Core standards. However, neither teacher recognized these goals as worthwhile in light of the demands of the next day's instruction.

In other cases, teachers agreed with the goals of PD sessions, but were frustrated by their execution. For example, Diane expressed frustration in year 4 with an online workshop about African American students and the achievement gap. "[It is] not really support, it's been mainly

like ‘these kids can’t do this because they come from bad home lifes’ [*sic*] and nobody ever offers you any suggestions about how to fix it” (year 4 interview).

In District B, many teachers were unenthusiastic about attending the Stipend Sessions that they were compensated for attending during the year. Mario, an improver, noted repeatedly in years 5-7 that he didn’t enjoy Stipend Sessions because of the general ‘atmosphere’ created by the negative attitudes of many of teachers who attended. “A lot of [teachers]...they don’t care to listen, and a lot of the people who lead it are also...I don’t know, it doesn’t seem like they care too much either. So, it’s kind of like, everybody’s forced to do this when they don’t want to be there...” (Mario, year 5 interview). Anna had similar frustrations with the Stipend Sessions, saying “it feels like a lot of people are complaining, you know ‘well we don’t have this, we don’t’ have that...’” (Anna, year 2 interview). Even when the session conversations focused on particular upcoming tasks, Anna struggled that first year to apply what she learned or talked about during the sessions, particularly because her school’s class periods were half as long as those at many of the other participating teachers’ schools. With less time for each task, she and the other 6<sup>th</sup> grade teacher at Park Falls often fell behind the district’s pacing calendar – a problem that she found could not be solved with the “little hints [about] ‘here’s what you can do to help catch up’” that were shared at the sessions (Anna, year 2 interview).

This set of findings, though limited due to the nature of interview data, indicate that formal PD may provide opportunities for teachers to experience and perhaps develop a vision of more ambitious instruction through participating in inquiry-oriented math as students might. Diane and Anna both described the summer CMP institutes they attended as helpful for just this reason. However, given that nearly all of the stable-procedural teachers and decliners also reported attending some math-specific PD that emphasized an ambitious vision of instruction, in

some cases the same exact sessions that improvers attended, it seems that these sessions were not, on their own, sufficient for supporting teachers to develop more ambitious practice themselves.

**Teacher Collaborative time (TCT).** The third form of support I analyzed was school-based TCT that teachers regularly attended. Using case teachers' interview reports, I categorized the kinds of activities that teachers regularly participated in during these meetings; I drew on school summary forms to triangulate these reports with those of other teachers in their TCT group. Using the larger MIST dataset, I also assessed the instructional expertise of the facilitator of these TCT meetings, when there was one, using that teacher or coach's IQA, VHQMI, MKT, and VSMC scores, when available.

My analysis of improvers' descriptions of their TCT meetings showed that, when compared with decliners, improvers were more likely to attend TCT that was led by someone with significant expertise in math instruction – i.e., a math coach or teacher with high VHQMI and/or high IQA scores. Decliners, by contrast, usually attended TCT meetings with no facilitator (e.g., teachers would rotate responsibility) or a facilitator with no more expertise than the teachers in the room. However, despite having TCT leaders with more instructional expertise, I did not find any clear differences in the kinds of activities in which the teachers in each group engaged during TCT. Instead, I found that across all case teachers (improvers, decliners, and stable teachers) TCT meetings in most schools were usually spent mapping out pacing or materials for upcoming weeks, or analyzing student data in order to plan interventions or group students into tutorials. These activities, while useful for helping teachers align their pacing schedules with that of their colleagues, generally did not involve discussions of specific teaching

practices or their affordances for student learning. In some cases, teachers did say that they would use TCT to share “what worked and what didn’t” in past or upcoming lessons (e.g., if someone was ahead of the other teachers). Cathy, who usually found herself several days ahead of the rest of the 8th-grade teachers, described her TCT meeting as an opportunity for her to brief other teachers about investigations that she had found problematic with her students, and what adjustments or supplemental materials she recommended. Other teachers described similar conversations, particularly in conjunction with looking at student assessment outcomes and achievement data. For example, many teachers described looking at student scores with other teachers from their grade levels, saying that they would look to see whose students did the best and ask what they did differently. In most cases, teachers did not describe actually working through individual lessons or investigations, or discussing particular practices like grouping or questioning strategies they would use in upcoming lessons; rather, their discussions of teaching practice seemed to more closely resemble the “targeted re-teaching” or “tips and tricks” forms of data-based discussions as described by Brasel, Garner, Kane, & Horn (2015). According to Brasel et al, *tips and tricks* discussions are often limited to sharing “superficial, one-size-fits-all strategies that do not adequately address the source of students' learning difficulties” (Brasel et al, 2015, p. 2).

This prevalence of low-level activities in teachers’ TCT meetings reflects the districts’ limited focus on and support for TCT in the early years of the study, beyond communicating expectations to principals that they should provide regularly-scheduled time in the school day for teachers to meet. Until years 5-7, districts B and D had few expectations for what teachers should be doing during this time, how instructional coaches should be involved, and how these meetings would support teachers’ development. Three of the improvers’ data came from years 1-3, when

most of the support for coaches focused explicitly on the CMP materials (e.g., PD that was similar/equivalent to what teachers were getting), or on one-on-one coaching techniques (e.g., cognitive coaching or the coaching cycle). Neither Karon (Diane and Cathy's coach), nor Lisa (the TCT leader at Anna's school) reported receiving any support or targeted PD to help them facilitate TCT.

Mario was the only improver whose data came from the later years of the study (y5-7). In the summer prior to year 7, District B math leaders and members of the MIST research team collaboratively developed a lesson planning protocol that teachers could use to guide TCT meetings that were of greater depth. After collaboratively planning their lesson, the protocol called for the teachers to reconvene to analyze student work and discuss ways they might improve upon their lesson design. Two Lagoons adopted this protocol for its grade-level planning meetings, though there were differences between the grade-level groups in the degree to which they found it a productive use of their time. Mario and the other 6<sup>th</sup> grade teachers used the protocol as their 'default' meeting agenda, but did not comment on whether or how this protocol helped him design or implement better lessons.

Lois, the stable, procedural teacher at Mario's school, also described the use of this lesson planning protocol in her 7<sup>th</sup> grade team meetings. Lois, however, was outspoken about her disdain for it. She found it a waste of time to spend an entire meeting planning a single lesson – a complaint that she also had about many district PD sessions that focused extensively on one aspect of instruction (e.g., the launch) or one particular lesson/investigation. In her view, these meetings and PD sessions wasted too much time addressing problems that she did not consider significant (e.g., planning launches or lessons in order to anticipate/support student reasoning).

Three of the decliners (Penny, Li, and Alishia) did describe engaging in some potentially high-leverage TCT activities at some point during their focal years, though all three also describe spending most meetings pacing out upcoming weeks' instruction. In Penny's case, these more high-leverage activities occurred in year 1; her coach, who had significant math instructional expertise, led TCT meetings in which teachers modeled practices or entire CMP investigations for one another. However, this coach left after year 1, and TCT at Meadowbrook in years 2-3 primarily consisted of using quantitative district assessment data to group students for pull-out tutorials. Thus, while the opportunities to engage in collaborative enactments may have provided productive opportunities for Penny and others to learn, these opportunities diminished quickly after the relatively expert facilitator left for a district-level position.

Li and Alishia (at Salt Marsh and Cypress, respectively) also describes some potentially high-leverage activities in TCT as well, such as planning lessons and looking at student work with other teachers in their grade level groups. Both of these teachers' grade-level groups did have more expert teachers in them, though neither of these TCT groups had an assigned facilitator who was positioned as a support for other teachers' learning. It is possible that these activities included discussions that explored the "why and how" of supporting student learning through particular forms of instruction (Brasel et al, 2015) and may have provided opportunity for these and other teachers to develop more ambitious vision or practice. However, in Li and Alishia's cases, it seems that these activities were not sufficient for them to maintain the ambitious forms of practice that once characterized their instruction. It is important to note that both Li and Alishia had quite unproductive views of their own students' mathematical capabilities; it may have been that these unproductive views shaped their engagement in these TCT meetings, or the extent to which they felt they could apply what was being discussed to

their own classrooms. I return to this conjecture in the next major section, in which I present findings related to my second research question.

In summary, my analysis of teachers' descriptions of their collaborative meetings found few clear distinctions between improvers and decliners that would suggest that TCT was a central support for improvers' development of more ambitious practice. These findings suggest that more research is needed which investigates what TCT leaders need to know and be able to do to productively support teacher learning. I return to these and further implications in the final chapter of my dissertation.

**School leaders' Instructional Expectations and Vision.** My analysis of teachers' instructional leaders focused primarily on teachers' perceptions of what the instructional leaders expected of them in order to be effective math teachers. My primary goal was to explore the conjecture that teachers are more likely to develop ambitious instruction if they are both supported *and held accountable* for doing so. While the scope of my research questions did not call for extensive analysis of their school leaders' expertise or practices, I did examine the VHQMI of each teacher's school leaders.

Teachers' descriptions of their school leaders' instructional expectations did not vary systematically between groups (e.g., improvers vs. decliners vs. stable teachers). In fact, there was very little variation across cases in what teachers perceived themselves to be accountable to. According to most of the teachers, their principals wanted them to have high levels of student participation/engagement during lessons, post daily lesson objectives, keep student work posted in the classroom, and follow the district's pacing guides. Some teachers mentioned that their principals looked for them to be using higher-level questions (often in reference to Bloom's

taxonomy), and in some cases, checked whether they were using CMP in daily classroom instruction. However, even in these cases, teachers generally perceived themselves to be free to use CMP as they saw fit. For example, Roger (the stable, procedural teacher at Diane's and Cathy's school) and Lois (the stable, procedural teacher at Mario's school) both said that their school leaders expected them to follow the district curriculum guides and use CMP for daily math instruction. To comply with these expectations, Roger adapted CMP into his very traditional lesson structure (demonstration, guided practice, independent practice), and Lois supplemented the curriculum heavily with other worksheets, since she was convinced that CMP was intended by its original authors to be a secondary resource that should be paired with a more traditional textbook. Knowing that her principal expected her to use CMP as the primary text, Lois used stickers to cover the publisher/resource name at the bottom of the worksheet before making copies.

These findings concerning leaders' instructional expectations are somewhat surprising considering that there *was* variation in the sophistication of school leaders' VHQMI. For example, the principal at Anna's (improver) school had a fairly traditional vision of instruction all three years during which Anna's instruction was improving. In particular, Anna's principal focused on regular and systematic analysis of quantitative state or district assessment data for selection/placement of students in various forms of tutorials happening before, during, and after the school day. In particular, his goal was to put systems in place that would quickly raise the achievement of their lowest-performing group, African American students. The result was a deliberate class assignment that grouped high-performing students together, separated low-performing African American students into a single classroom, and put the rest of their low-performing students into a third group, consisting largely of Hispanic students. In Anna's case,

her teaching became more ambitious despite her principal's expectations for daily classroom instruction, and his surprising and troubling approaches to improving the achievement of low-performing students. Given Anna's role as a coach and the time she spent with district math leaders, it may be that she came to identify more strongly with the district math leaders' vision of improvement than that of her principal, despite being tasked with the planning and data preparation needed to facilitate his approaches.

By contrast, two of the decliners (Alishia and Penny) had school principals who had quite sophisticated visions of math instruction. However, neither of these teachers perceived themselves accountable to implementing any particular forms of instruction. It is also worth noting that these teachers had both been teaching for many years, and while they had quite sophisticated visions of inquiry-oriented instruction, they also had quite unproductive views of their own students' potential for engaging in rigorous mathematics. In other words, Penny, Alishia, and their school leaders were all in agreement about what constitutes high quality math instruction; however, Penny and Alishia did not view their own students as capable of engaging in those forms of instruction because of behavior problems they encountered or deficits they ascribed to students.

In summary, though there were several cases of principals who described quite sophisticated, inquiry-oriented visions of instruction, there were no clear differences between improvers and decliners in what teachers perceived themselves accountable to in their instruction. Furthermore, even in cases where stable, procedurally-oriented teachers did perceive themselves accountable for using the CMP curriculum, their school leaders did not communicate any particular expectations for how these lessons should be implemented. These teachers' lack of

identification with an inquiry approach to math led them to adapt or supplement the curriculum in ways that significantly reduced their cognitive demand.

## **Part 2: Instructional Expertise, Identification with Reform Practices, and Instructional Improvement**

In this section, I present findings related to my second research question: *how might teachers’ instructional expertise and their personal identification with their district reform efforts mediate their engagement with available supports for their development?* To answer this question, I looked for potentially critical differences between the expertise and professional identification of improvers and their stable, procedurally-oriented colleagues who worked in the same schools and had access to the same school-based supports. For reference, the improvers and their stable counterparts are listed below in table 7.

Table 7.						
<i>Case Summaries: school, case type, years, and grade levels.</i>						
School	Case type	Name	Years exp	MIST years	District	Grade level
Cypress	Improver	Diane	5	2-4	D	6
	Improver	Cathy	6	1-4	D	8
	Stable	Roger	4	1-2	D	6
Park Falls	Improver	Anna	4 (1 in math)	2-4	B	8
	Stable	Roxanne	3 (1 in math)	2-3	B	8
Two Lagoons	Improver	Mario	2	5-7	B	6
	Stable	Lois	6	5-7	B	7

I found that three of the four improvers (Diane, Cathy, and Mario), unlike their stable, procedurally-oriented counterparts *and* all four decliners, came to identify with their districts’ reform efforts than their procedurally-oriented counterparts. My findings suggest that potentially necessary conditions for this sense of identification are (1) an inquiry-oriented visions of

instruction and (2) the view that their own diverse students (even those who were struggling) were capable of engaging in rigorous, inquiry-oriented mathematics.

In the remainder of this section, I elaborate these findings by stepping through each pairing of improver(s) and their stable, procedurally-oriented colleagues, describing similarities and differences in their instructional expertise, the degree to which they identified with the district reform efforts, and differences in their individual engagement with the available supports for their learning.

**Diane and Cathy (improvers) and Roger (stable) at Cypress Middle School.** My analysis of these two improvers and their stable, procedurally-oriented colleague found that all three teachers had numerous opportunities to access support for implementing CMP2 (PD, coaching, school-sponsored PD, opportunities to observe other teachers). However they differed greatly in the forms of teaching with which they identified, and this difference seemed to impact their views of the district's reform efforts, of the usefulness of the CMP curriculum, and their interest in getting support from school- and district-sponsored PD and coaching.

***Diane and Cathy: vision and identification.*** Cathy and Diane (3<sup>rd</sup> and 6<sup>th</sup> year teachers, respectively) both had very sophisticated, inquiry-oriented visions of math instruction, and held themselves accountable to that standard with all of their students. For example, though Cathy had been using CMP2 for several years at her previous school, the 8<sup>th</sup> grade teachers at Cypress would not adopt CMP2 until the following year. This presented a number of challenges to Cathy. She found it very difficult to assimilate into the normative teaching practices of the other 8<sup>th</sup> grade teachers, whose classrooms were characterized by quiet students doing independent practice after a demonstration by the teacher. "My classroom doesn't look the same, where the

other teachers sit and the kids are, you know, at their seat and all working quietly. I hate that...I can't stand it" (Cathy, year 1 interview). Though she knew it set her apart from the others, Cathy maintained that student collaborative work on more conceptually oriented tasks, rather than teacher demonstration and independent practice, were best for her students. "That's not what we came here to do because...I feel like we learn from each other, and I definitely prefer group work and partner work" (Cathy, year 1 interview). She added that this view stemmed from her experience as a CMP2 teacher:

There are a lot of teachers in our district that are not a fan of Connected Math because they feel it doesn't have enough practice...and I think that's really based on what it is you're used to teaching...they had taught from the [traditional] book, so when they had to go to this abstract way of thinking, it was hard for them. So, I came in having to do Connected Math, and never having a math textbook, so it was hard for me to go to the [traditional books at Cypress]. But I would probably prefer the Connected Math (Cathy, year 1 interview).

Diane similarly identified herself as "on board" (year 1 interview) with CMP2, realizing that it was not a popular choice among many teachers and parents because of the alternative vision of instruction guiding the creation of the materials. "The parents call and they're like 'where are the examples in the book?' and you don't have those, because [CMP2 involves] throwing a problem at the kid, and you're not giving, you're not standing up there teaching and they're not taking notes and stuff."

Though they both regarded CMP2 and rigorous, inquiry-oriented instruction more broadly as the best way to support students to be successful in mathematics, they both faced the same challenges that their colleagues did – chief among them, students who came to their

classrooms lacking basic skills and number sense. In interviews with MIST researchers each year, both Diane and Cathy described this challenge as a persistent, an often perplexing one. “I can teach my pretty little heart out, and they can get the concepts and they understand, and they got it, and then they mess up on addition or subtraction or multiplication, and it just breaks my heart....they don’t have reasoning skills....so we’ve been trying to work on that this year. [We talk about why] that number would be too big and other stuff. But, they don’t do that on their own, that’s not a skill that they have. That’s the answer that the calculator gave them so that’s what they wrote down” (Cathy, year 1 interview). Diane described similar challenges with building her students’ number sense in her year 3, and the difficulty entailed in building students’ reasoning skills over time. “It’s like they’ve never had to [use number sense]. And so I think [CMP2] is helping because...it’s trying to get them to think. It’s just the teachers have to keep pushing and pushing, and waiting, and it’s frustrating for the teachers and it’s frustrating for the kids.” (Diane, year 3 interview).

Despite these difficulties, both Diane and Cathy continued to frame these challenge of student learning difficulties – even those that were most exasperating to them – as challenges that could be addressed through additional supports and improved instruction. For example, Cathy used group work frequently so that students could use one another as resources in solving problems. She also made calculators available to her students in every class, and provided students with extra time to work on investigations at home if they needed it (year 4 interview).

As another example, the school schedule at Cypress beginning in year 3 included an additional period that rotated between math and science enrichment. Since the teachers were given no set curriculum for this extra time with their students, Diane and the other 6<sup>th</sup> grade math decided to use the period to preview/review the basic skills their students would need for

upcoming CMP2 investigations. This use of the secondary math class to provide basic skills supports to students *in anticipation* of the difficulties they might face was markedly different from the approach of many schools in the MIST sample (including Penny's, Li's, and Anna's schools, and even at Cypress in years 5-7). More commonly, these extra periods were used for test prep or skills review that was unrelated to the topics or investigations students were doing in their primary math classes.

For both Cathy and Diane, their goal in providing supports to struggling students was to help them be successful in rigorous, inquiry instruction. However, despite having ambitious goals for all their students, and identifying strongly with an inquiry-oriented vision of instruction, both encountered cases of students who refused to engage in class activities, or for whom they had exhausted all other strategies and ultimately resorted to guiding them through problems step-by-step. However, as Diane described, these cases were exceptions rather than the rule:

[While some students work on enrichment activities], this side of the room is gonna work with me on...this, this, and this. And we're **gonna go back through** this investigation and, you know, sometimes as much as I would love to let them discover formulas and stuff on their own, sometimes with certain kids you just have to tell them 'these are the steps that you take, and then we're going to practice them'...because they need that...it takes them so long to get it that they end up shutting down if you don't help them along at some point (Diane, year 3 interview; emphasis mine).

It is important to note how in the above example, Diane describes using these differentiated supports *after* using the same problem-solving task with the entire class, "going back through it" with them while other students did computer-guided enrichment activities on the

same concept. Rather than substituting rigorous tasks and student-led inquiry with guided practice and step-by-step instruction, both Cathy and Diane looked for ways to support struggling students that preserved their high expectations for all students. They both saw CMP2 as the best tool to accomplish *their* goals for all students in the primary math classroom, and looked to more expert others for support when they ran out of ideas for how to improve.

*Diane and Cathy: accessing support to improve.* As teachers who identified with the district’s vision of ambitious instruction, and yet continued to face daily challenges with students who struggled in rigorous math instruction, Cathy and Diane were both motivated to seek support from others with more experience and expertise using CMP2. For example, Diane and Cathy both volunteered to be part of a small cohort of teachers who would receive additional support from district leaders as they implemented the new curriculum. These sessions occurred throughout the school year and mirrored the summer PD institutes, and often included time to work through, plan, and rehearse upcoming investigations or entire units with support from district or contracted coaches who had extensive experience with the CMP2 materials. Their participation in this cohort also gave them opportunities to observe more experienced teachers at other schools teach CMP2 investigations, and positioned them in their schools as experts for other teachers to observe. Diane described looking forward to these sessions, particularly because they gave her opportunities to ask for advice on “how to modify something, [and about] what doesn’t work” (year 1 interview).

In year 2 (Diane’s first year at Cypress, and Cathy’s second), their principal used school funds to provide supports for the math department to implement CMP2. First, he hired Karon, a former high school math teacher, to be a math coach whose primary responsibility was to support teachers’ implementation of the new curriculum. In year 2, both Diane and Cathy worked

regularly with Karon, who regularly visited their classrooms to observe, co-teach, and work through tasks with them as she herself learned more about the curriculum. Though Karon focused more on other teachers in subsequent years, both Cathy and Diane continued to work with her for at least a week each semester when she rotated through their classrooms to offer modeling and feedback on their instruction.

Secondly, in February of year 2, the principal funded a trip for all members of the math department and an assistant principal to a three-day CMP workshop held at Michigan State University. Both Cathy and Diane looked forward to the workshop, seeing it as an opportunity to ask questions and get support. For example, Cathy, was concerned at the time about how to improve her classroom management skills as she worked with struggling students during CMP2 investigations: “I don’t know how to fix it, so I’m hoping when I go to this CMP [workshop] that they’re going to help me fix it” (Cathy, year 2 interview). Diane, meanwhile, was hopeful that the weekend would help change the minds of others in the school who were less enthused about CMP and their principal’s requirement that they all use the materials for daily math instruction. Roger was one such teacher.

***Roger: vision and identification.*** Roger, who three years of teaching experience prior to year 1 of MIST, was aware that his district, school leaders, and his math coach all had similar goals for improving math instruction. However, unlike Cathy and Diane, Roger had serious objections to these goals. First, he saw his school’s adoption of CMP2 as a politically-driven act of compliance with a district decision that impinged on his and other teachers’ autonomy. As such, he felt that CMP2 (and the requirement that he use it) represented a threat to democracy and autonomy at the school level, and to the principles of site-based decision making more broadly. He described in years 1 and 2 how he actively resisted the districts’ reform efforts by

speaking out against CMP to his colleagues and school leaders during TCT meetings and in private conversations, and by attempting to rally the other math teachers to petition the district to vote the curriculum out.

Secondly, Roger also disagreed with many of the teaching practices he saw represented in the curriculum, such as student-led inquiry, group work, and extended time for mathematical discussions. “You know, I don’t think good instruction is just to say, ‘oh here group, read through this investigation and figure it out for yourself.’ You know I think the students need more guidance than that” (Roger, year 2 interview). When describing the role of discussion in math, Roger said –

[Discussion] can be helpful [...] to see if the kids had any valid points to bring to the table [...] but most of the time the teacher teaches it and the students kind of take it in. I mean if they have...any kind of question about it, then they should feel free to ask. But I mean, there’s not a lot of room for debate on the math, because you know, this is it (Roger, year 2 interview).

Facilitating this kind of instruction (teacher demonstration followed by independent student practice) using CMP2 proved quite difficult for Roger. Since CMP2 is structured differently than traditional textbooks, with worked examples followed by practice problems, realizing his own vision of instruction with his students entailed walking the entire class through CMP2 investigations step-by-step, or using other resources to present materials. In year 1, for example, Roger reported getting negative feedback following an observation by his principal, who wanted to know why he was not using CMP2 on the day he was observed. “It’s because we were already several weeks behind where we needed to be, and I was trying to combine some stuff and just kind of teach it my way to try and get us caught up” (Roger, year 1 interview). Both

years he taught math at Cypress, Roger continued to contrast “his way” with CMP2, including his stance on group work. “I don’t use groups probably as much as I should. But I don’t think that groups are always needed. I think students need to do their own work most of the time” (Roger, year 2 interview). This gap between the district’s improvement goals and Roger’s own vision for instruction is especially pronounced when contrasted Diane and Cathy, who saw CMP’s emphasis on student collaboration and discovery as affordances rather than nuisances

In conjunction with his very traditional VHQMI, Roger also had quite clear views on what led to differential outcomes among his students: namely, students’ self-discipline or motivation to stay engaged. In Roger’s view, learning mathematics was a simple matter of paying attention. “Usually the teacher’s going to teach it in a correct way. They’re not usually going to give bad information, so if everybody’s paying attention then it should get across to them as long as it’s explained very well or shown” (year 2 interview). Multiple times in his interviews during years 1 and 2, Roger reiterated this framing of mathematics teaching and learning as a one-directional giving and receiving of knowledge, and unsurprisingly used a similar framing when describing the learning outcomes of African American students – a sub-population which was underperforming at Cypress in year 2.

Roger: I don’t look at, you know, [students’] ethnicity. You know, I see that, are you paying attention? Are you doing what you’re supposed to do? Then you’re probably going to get what you need to get. If you’re disrupting class, then you’re holding everybody back.

Interviewer: Have you received any professional development targeted specifically to helping low-performing students?

Roger: ...They've had some [PD days] here. We talked about how to, how to interact with black students in particular.

Interviewer: Okay; have those – were those helpful?

Roger: I would say no because I already, I already know how to interact. You know it's, it's like I said – it's not the race that I look at. It's 'are you doing what you need to be doing?' You know, if you do, then we'll be ok.

Roger also saw clear differences between his regular classes and his advanced classes – namely, that with the latter he was “able to let them try and work things out more on their own, try to figure stuff out for themselves more than with [my regular classes]. They can think like that and they're motivated to do that” (Roger, year 1 interview). By contrast, with his regular math classes, he would “read through it all and work through it together...” (Roger, year 1 interview) rather than letting them read and reason through the investigations on their own. In contrast to Diane and Cathy, who resorted to procedural instruction for exceptional cases after exhausting other options, Roger's VHQMI and VSMC indicate that he not only saw traditional, procedural instruction as high-quality, but that he also saw reason to reduce the rigor of that instruction even further for students who were not in advanced classes.

***Roger: engaging with supports for improvement.*** In the summer prior to MIST year 1, Roger attended a multi-day CMP institute held at the district PD center that introduced him to the CMP materials. He said this was helpful for getting 'acquainted' with them, because otherwise he would not have known how to use them. He did not attend any other PD during the school year of MIST year 1, saying he was too busy to attend them due to his role as the coach of the 'Quick Recall' academic team, in which students competed with other teams to answer trivia questions (including but not limited to math) as quickly and accurately as possible. This coaching also kept

him from attending most department meetings. Thus, the only PD Roger attended in year 1 was that summer institute; unlike Diane and Cathy, he did not volunteer to be part of the 6<sup>th</sup> grade CMP cohort. The principal at Cypress did offer some additional supports during year 1; for example, he offered to pay for substitutes to allow Roger to observe other teachers who were more proficient with CMP. However, Roger was not very interested. “I feel like it’s better for me to be in here trying to move us forward than to be in another class observing and just having a sub in here. I think that would be more help” (Roger, year 1 interview). Given Roger’s resistance to the curriculum and the forms of instruction he saw represented therein, it is not surprising that Roger saw little point in observing other teachers’ classrooms – especially in year 1, when he was still sure that he and the other math teachers would be able to vote the curriculum out at the end of the year.

In year 2, school-based supports for the math teachers at Cypress included the math coach (Karon), a school-funded trip for the math department to attend the weekend CMP2 workshop at MSU, and ongoing opportunities for math teachers to observe one another’s classrooms – with Karon substituting for them while they were out of their classrooms. Unlike Diane and Cathy, whose challenges in implementing CMP2 seemed to motivate their engagement with these supports (e.g., Cathy looking forward to the trip to MSU as an opportunity to “fix” her problems with classroom management), Roger was unenthused, and at times skeptical, about these same supports.

In year 2, at his principal’s insistence, Roger did observe both Diane and Cathy teach one lesson each. He described observing Diane as helpful because she taught the same grade level (whereas he said he did not get much from watching Cathy teach an 8<sup>th</sup> grade lesson). He did not say what he learned from Diane’s classroom, only that he picked up a couple of ideas. Whatever

he saw in her lesson, however, his vision of high-quality instruction and his description of how he led students through CMP tasks step by step to demonstrate procedures both indicate that these observations did not substantially change his views. Karon, the math coach, commented on the mixed outcomes of this effort to have CMP-resisting teachers observe others in the school:

I think that's been...probably been the most beneficial because when the teacher comes, you know, leaves [the other classroom], they are like, 'wow, those students did really well with it,' and you know, they are asking questions and they were really impressed with the classroom. But then what amazed me on the flip side was, 'Well, those aren't my kids, my kids aren't like that.' And I thought, how could you say that? I mean...it's not like that teacher has any better student than what you have. Every teacher in this building ...has [the same course load], so I wouldn't think that their [regular] classes are any better than your [regular] classes [...]. So I think that maybe that is what they tell themselves to keep on doing what they are doing (Karon, year 2 interview).

Given Roger's unproductive views of the mathematical capabilities of students in his regular mathematics classes, Karon's conclusions are likely correct: Roger's observations of other, more inquiry-oriented classrooms were likely framed (and thus constrained) by his views of his own students' capabilities. Because he did not regard his own 'regular' students as capable of doing the thinking and reasoning called for by CMP2 investigations, he likely did not view the practices he observed in Diane's classroom as particularly applicable to his own students.

Perhaps because of his views his own students' capabilities, Karon made a number of attempts to model instruction in Roger's classroom. "Frequently she'll volunteer to teach," Roger said in year 2. "She'll say, 'Oh I'm doing this lesson with this other teacher, do you want me to come in and do that with you?' And so I think she might be doing one before too long in here"

(Roger, year 2 interview). However, by the spring, when Roger completed the MIST teacher survey, he reported only observing Karon teach a lesson “once or twice.” Furthermore, it is possible that even when observing Karon teaching his own students, Roger did not see her as a source of expertise or support for improving his instruction. In Roger’s view, Karon’s was a “liaison between the administration and the teachers” (Roger, year 2 interview) who would communicate any information from the administrators during department meetings. While he said that she was “constantly” in his classroom in year 2, he viewed her as there to provide “backup” and “help any students that have questions” (Roger, year 2 interview). He recognized that she was there to help him implement CMP, and that she wanted him to use group work more; however, he did not seem to view her as a source of expertise or someone to learn from, but rather someone whose presence in his classroom was more evaluative than supportive (see fig 1, an excerpt from Roger’s year 2 MIST survey, below). “If she has something say...something to add, then she will. And you know, she doesn’t you know, say things, add things too much...” (Roger, year 2 interview).

<p>To what extent do you agree or disagree with each of the following statements?</p> <p>The purpose of the mathematics coach visiting teacher's classroom is to directly assist teacher in improving my teaching</p>	<p> <input type="radio"/> Strongly disagree  <input type="radio"/> Disagree  <input checked="" type="radio"/> Neither disagree nor agree  <input type="radio"/> Agree  <input type="radio"/> Strongly agree  <input type="radio"/> N/A         </p>
<p>To what extent do you agree or disagree with each of the following statements?</p> <p>The purpose of the mathematics coach visiting my classroom is to evaluate my teaching in terms of job performance</p>	<p> <input type="radio"/> Strongly disagree  <input type="radio"/> Disagree  <input type="radio"/> Neither disagree nor agree  <input checked="" type="radio"/> Agree  <input type="radio"/> Strongly agree  <input type="radio"/> N/A         </p>

Figure 1. Roger’s responses to coach questions on his year 2 survey.

In February of year2, Roger attended the school-funded trip to the MSU with the rest of the math department. Unfortunately, we could not question about this PD because he left the study to become a social studies teacher at Cypress. However, he was asked about the (then upcoming) trip during his year 2 interview. He responded with strong skepticism, saying that he was “open-minded...kind of” (Roger, year 2). “I have a lot of questions for them” such as how they expected students with low reading abilities to participate in CMP2, or why certain mathematical concepts were presented in ways he disagreed with. His responses to the MIST teacher survey (which was completed after the MSU trip) indicate that he maintained very negative opinions of both CMP2 and the PD he had attended that year (see figure 2 below).

<p>To what extent do you agree or disagree with the following statements about school and district professional development sessions this school year?</p> <p>Advocated practices I do not believe in</p>	<p><input type="radio"/> Strongly disagree</p> <p><input type="radio"/> Disagree</p> <p><input type="radio"/> Neither disagree nor agree</p> <p><input type="radio"/> Agree</p> <p><input checked="" type="radio"/> Strongly agree</p> <p><input type="radio"/> N/A</p>
<p>To what extent do you agree or disagree with the following statements about school and district professional development sessions this school year?</p> <p>Made me question my beliefs and assumptions about which teaching methods work best with students</p>	<p><input checked="" type="radio"/> Strongly disagree</p> <p><input type="radio"/> Disagree</p> <p><input type="radio"/> Neither disagree nor agree</p> <p><input type="radio"/> Agree</p> <p><input type="radio"/> Strongly agree</p> <p><input type="radio"/> N/A</p>
<p>To what extent do you agree or disagree with the following statements about school and district professional development sessions this school year?</p> <p>Were consistent with my own goals for instruction</p>	<p><input type="radio"/> Strongly disagree</p> <p><input checked="" type="radio"/> Disagree</p> <p><input type="radio"/> Neither disagree nor agree</p> <p><input type="radio"/> Agree</p> <p><input type="radio"/> Strongly agree</p> <p><input type="radio"/> N/A</p>

Figure 2. Roger’s survey responses pertaining to his views on CMP2 (year 2).

**Summary.** The contrasting cases of Diane, Cathy, and Roger illustrate relationships between teachers' individual identification and goals for instruction – and how this identification may shape teachers' engagement with professional development supports. Unlike Diane and Cathy, who identified with the principles of teaching and learning they saw represented in CMP2, Roger viewed the introduction of CMP2 as the result of a political whim, and as a threat to his own autonomy and values. He viewed math instruction as a straightforward process of giving and receiving, where instructional quality can be gauged by correctness of the information the teacher imparts, and by students' attentiveness. Roger never came to problematize his practices or his views of math and student learning, despite having opportunity to see his own students and students like them participate in inquiry-oriented instruction.

Unlike Diane and Cathy, whose difficulties with particularly challenging students motivated them to seek out support from others with more experience and expertise (e.g., joining the district CMP cohorts, and seeking out advice from their coach and at MSU institutes), Roger seemed to take his students' difficulties with CMP2 as proof that they were incapable of the kind of thinking and reasoning the curriculum called for – particularly students in his 'regular' classes. Roger attended very little CMP2-affiliated PD during his two years teaching math at Cypress, and had very negative opinions of the PD that he did attend.

After year 2, Roger stopped teaching math at Cypress, and went back to teaching social studies. It is unclear whether this decision was his own, or whether his principal had a hand in this move.

**Mario (improver) and Lois (stable) at Two Lagoons.** In my analysis of the contrasting cases of Mario and Lois at Two Lagoons Middle School, I found that both Mario and Lois had

problems implementing CMP2 with their predominantly Latino, ELL students. However, they framed these challenges quite differently. In particular, Mario's and Lois's contrasting views of their students' capabilities, and how to support struggling students, resulted in substantial differences in their instruction. Similar to Diane and Cathy, Mario's difficulties in reaching particular students seemed to motivate him to seek out support from others with more expertise. Lois, by contrast, used her students' prior histories of failure to justify her use of low-level instructional activities – even when it required skirting her administrators' expectations that she use cognitively demanding tasks.

***Mario: identification and vision.*** Unlike Diane and Cathy, Mario's trajectory of instructional improvement is not linear. Though Mario's improvement occurred in years 5-7, his first year of instruction (year 4) and his experience implementing CMP2 as a brand new teacher are relevant to the years that followed.

Mario came to Two Lagoons MS as a new teacher in MIST year 4, and was energetic and positive about his school, his administrators, the school-based supports he was receiving, and the inquiry approach that CMP used to support student learning in mathematics. Similar to Cathy and Diane, Mario's vision of high quality math instruction aligned with what he saw represented in the CMP2 curricula. In particular, Mario appreciated how CMP2 presented problems that were “tough” and provided students opportunities to develop “conceptual understanding that [would support them to] answer that problem down the road” (Mario, year 4 interview), not just items on the year-end state test. He maintained this view of CMP2 in subsequent years, saying in year 6 that it “asks really high level questions...has students working together...[and] relates it to real life” (Mario, year 6 interview).

Despite recognizing these positive qualities of CMP2, Mario became very critical of CMP2 over years 4-7 – in particular, in whether he could adequately meet the learning needs of his students by doing “each question by the book” (Mario, year 6 interview). He described the difficulty of explaining task scenarios such as *Bike Touring in New England* to his students, and wished for a curriculum that could had CMP’s focus on conceptual understanding but featured task scenarios that were more familiar to his students. He also struggled with how to balance ambitious goals for developing his students’ problem-solving skills with the need to prepare students for procedurally-oriented state tests. “In a perfect world we wouldn’t have tests or [skills-oriented] problems for them to do, but we do have it...” (year 6 interview). Knowing that his students would ultimately be held accountable for the kinds of assessed by the state test, Mario also felt obligated to incorporate opportunities for them to practice those skills.

Mario’s approach to addressing these challenges evolved over time. In his first year, Mario frequently went to Vera, his assistant principal who was a former district math specialist, for advice in how to modify CMP tasks in ways that would maintain ambitious goals while keeping him up to date with the district’s pacing guide. In year 5, he went through what he later described as a “rebellious year,” and took a break from the “battle” of making CMP2 work for his students (year 6 interview). He instead opted to generate his own classroom tasks and activities using released state assessment items. Mario’s IQA scores from year 5 indicate that the rigor of his instruction dropped significantly as he focused on preparing students for the state test.

By the end of year 5, Mario was overwhelmed with the amount of work it took to fashion his own curriculum. And while he still saw CMP as a challenging resource that needed to be adapted for his students, he decided in year 6 that it was ultimately better, and easier, to modify

CMP than to start from scratch. Furthermore, despite its flaws, he felt that CMP was the best available tool for supporting students to develop mathematical “processing” or reasoning skills – a goal he seemed to neglect in year 5 when he focused exclusively on preparing students for the state test. During years 6 and 7 he returned to using CMP as his primary instructional resource, though he adapted CMP for his students by changing the task scenarios so that his students could make better sense of them. For example, Mario changed the names of people and places in task scenarios to nearby locations and familiar people or television characters so that they would require less explanation during the launch. As he became more familiar with his state’s standards, Mario also began to modify the numbers in CMP2 tasks so that his students made fewer computational errors that might derail opportunities to learn larger concepts. “I kind of look at the objective of what they’re trying to get them to learn or how hard it is and I kind of weigh it with ‘well what’s the big idea, what’s the objective of today that they have to know?’ and so I’ll change it” (Mario, year 6 interview).

Though Mario had a rocky relationship with CMP2 over years 4-7, he interpreted student learning difficulties in math very differently than did Roger. While Roger saw student learning difficulties as evidence that his students were not paying attention, Mario saw his students’ difficulties and lack of motivation as evidence that he was not meeting their needs adequately.

Like, [lack of motivation] could be parental involvement. It could be – I think the number one thing it could be is because I’m not doing a good enough job to motivate them. Like, a lot of teachers ignore that part, but it’s like I need to do a better job of giving them a reason to be motivated and having that buy-in and a hook at the beginning of the lesson (year 6 interview).

Similar to Diane and Cathy, Mario took responsibility for his students' learning – and struggled to know how to accomplish an ambitious vision of instruction with his students. Like Diane and Cathy, he saw student difficulties as evidence that he needed more support to do a better job. “I’m really burdened with the idea of like, how can I make sure that my students are processing [i.e., developing conceptual understanding] and that I’m assessing that and seeing that...I would really love to see somebody doing that successfully on a day-to-day basis” (Mario, year 7 interview).

***Mario: supports for improvement.*** Unlike the other improvers in this analysis, Mario did not have an opportunity to attend an extended CMP2 institute because he was not hired in time to attend the last of such institutes in District B. He did attend a number of professional development sessions during years 5-7, but there is no evidence that these sessions supported his development of more ambitious practice. He expressed mixed opinions about the helpfulness or relevance of these sessions, particularly pull-out Stipend Sessions that teachers were compensated for attending during the school year. For example, Mario found it helpful to attend PD sessions that included time to work on specific lessons or practices together with others – e.g., creating a launch for a particular CMP investigation, or bringing back student work and sharing what strategies teachers had found most effective with their students. However, he also said that many PD sessions left him feeling frustrated. For example, he said that the atmosphere at Stipend Sessions in years 5 and 6 ruined the experience for him: “even if we have a good presentation there’s teachers who don’t want to be here, sometimes the presenters don’t want to be here, and it’s kind of a drag” (year 5 interview). Surrounded by “people just not wanting to be there in general, or just older people who just complain and gripe a lot, or younger people

who just ignore stuff,” (year 7 interview) reduced the value of the sessions for him and other teachers who were seriously committed to improving their instruction.

Similar to the mixed opportunities to learn that Mario had in pull-out PD, TCT at his school seemed to offer limited support for instructional improvement during years 4-7 even though the district’s increased focus on teacher collaboration during those years had impacted the structure and leadership of those meetings to some degree. Mario was a part of the 6<sup>th</sup> grade TCT group at Two Lagoons, which had regular meeting time following each week’s department meeting. Though Mario’s opinion of TCT meetings did improve over years 4-7, it is unclear whether or how his participation in these meetings may have supported his instructional improvement. For example, the meetings with the entire math department were facilitated by the math department chair and part-time coach, Harry, who had a relatively sophisticated vision of instruction and was relatively accomplished in ambitious instructional practices. However, Harry was also regarded by Mario and the other math teachers as a poor facilitator and leader. In grade-level planning meetings with other 6<sup>th</sup> grade teachers, the primary activities Mario described were quite low-level, and did not involve opportunities to investigate or enact practice. Beginning in year 7, however, the 6<sup>th</sup> grade team regularly structured their meetings around the lesson planning protocol – a TCT facilitation protocol that was co-developed by district leaders and the MIST team. However, despite their use of this protocol, Mario remarked that his colleagues’ views on instruction often limited the helpfulness of their collaboration:

I try to often have conversations with the other math teachers about what they’re doing and how they’re doing it and what resources they’re using. Sometimes those conversations can be tough for me because I feel like from teacher to teacher, you just have a different perspective as to what’s high quality instruction.... I’ve got one, one guy

that I work with who is very adamant about having... a lot of basic computational practice... so he might plan a lesson about the same [standards] you know, 100% differently than me just because he's more willing to practice more, and I'm more willing to, you know, discuss it more (year 7 interview).

Mario's recognition of the difference between his own vision of instruction and that of his colleagues resembles Diane and Cathy, who also described having difficulties collaborating with others with whom they worked.

Mario's coaching support fluctuated during years 4-7 at Two Lagoons, in parallel with the quality of his instruction each year. In year 4, Mario went regularly to Vera, his assistant principal, who had formerly been a district math specialist. Mario saw Vera as a valuable source of expertise and advice for how to stay on pace with the district pacing guides while integrating the school's various initiatives into CMP2 investigations. He also got regular support in year 4 from Keisha, a district math specialist, who frequently observed him and gave him feedback on his instruction. Based on Mario's description of this feedback, Keisha was likely a critical support for his learning. For example, Mario said Keisha told him that he spent "too much time in front of the class lecturing...so I try to do the Kagan strategies, which is the small group work, but I feel like in those situations very often I lose students....so that's what I've been trying to figure out with Keisha here..." (Mario, year 4 interview).

The only year Mario did not receive any coaching support was in year 5, which he termed his "rebellious year" during which he abandoned CMP2 because he felt he "didn't know how to teach it in a way that felt effective" (year 6 interview). Given this lack of support from a coach with math expertise, it is perhaps not surprising that the tasks Mario developed on his own were low-level procedural tasks.

Beginning in year 6, Malcolm (a district math coach) was assigned to work with Two Lagoons among a number of other middle schools. Malcolm observed Mario a few times in year 6, but his visits were infrequent and short. Mario said he never heard anything back from those observations, even though he saw Malcolm taking notes while he was there. This changed in year 7; Mario said in his year 7 interview that Malcolm had been more present in the school and in his classroom, that they had co-planned lessons together, and that he regularly gave positive, but constructive feedback to him after observing his instruction. Malcolm also attended department and grade-level planning meetings about every other week in year 7, answering teachers' questions about upcoming units in the curriculum, and how they could catch up if they were behind in their pacing.

*Lois: vision and identification.* By year 5, when Mario was piecing together his own math curriculum for his 6<sup>th</sup> grade students, Lois had been teaching middle-school math for 6 years, middle-school social studies for 10 years previously, and high school math before then. She had been at Two Lagoons when the district first adopted CMP2, and she had expressed skepticism and frustration with the curriculum from the beginning. Lois's chief problems with CMP2 were twofold. First, similar to Roger, she felt that the district had been less than forthright in adopting CMP2 as the district-wide curriculum for middle school math. "I went to almost every single [text adoption forums] and of course we, or they, voted for [another publisher's text] and I couldn't agree more, and then we ended up going with [CMP2]" (Lois, year 1 interview). Furthermore, Lois felt (or perhaps had been told) that CMP2 was meant to accompany a more traditionally-organized textbook. "[CMP2] was designed to be in addition to another book by the same publisher, but our school district only adopted this one because they thought that teachers wouldn't use it if we had the other one, but so it's a problem" (Lois, year 5 interview).

Lois's second problem with CMP2 was that it conflicted with her own goals for instruction for her students at Two Lagoons. Her classes, the majority of whom were 'regular' 7<sup>th</sup> grade math classes during years 5-7, were composed of students who Lois described as scarred by prior failures – either in math classes, or on state tests that many of them had not pass. For these students, Lois felt that CMP2 was particularly inappropriate because it did not present mathematics through worked examples that students could refer to as they moved forward in the curriculum, but instead presented extended scenario-based investigations. “As you look through this book...it's like all words. I mean, if I was learning math, and I had to look at that, I would close the book, and I love math! ....[It] doesn't give the repetition that they need to master the skill” (Lois, year 5 interview).

Interestingly, when Lois described her vision of high quality math instruction during years 5-7, her interviews suggest that unlike Roger, she *did* value many aspects of inquiry-oriented instruction. For example, in year 6, Lois said high quality math instruction involved students being able to solve problems rather than just repeat procedures (“[students should be] having to problem solve. That it's not just like...a bunch of problems that look exactly the same”), and privileged conceptual classroom discourse (“they can defend their answers...there can be debate going on of why one strategy versus another, that they can teach each other”). Lois's development of a VHQMI which privileged problem-solving over procedures and conceptual discourse over rote drill may have been supported by her participation in a 2-year master's program offered by a local state university, in which she took one math methods class each semester beginning in year 6. Perhaps in part because of this program, or perhaps due to her history as a high school math teacher, Lois's MKT was also far more developed than most teachers in the district, especially her knowledge of algebra concepts.

Given Lois's relatively sophisticated VHQMI, it is at first surprising that she was opposed to using CMP2 as a primary resource for her students, and that her classroom instruction was consistently marked by a focus on procedures and drill across years 1-7. Based on her interviews in years 5-7, however, this gap between Lois's vision and practice is likely due to her view that her students, who were "many grade levels behind" (Lois, year 5 interview) in their mastery of basic skills, needed to experience success in lower-level problems and basic procedures in order to develop motivation to continue. During years 5-7, Lois continually framed her students' learning difficulties in mathematics as a problem of low motivation due to past failure – a problem that was only augmented by unreasonably high expectations that set them up for further disappointment. She felt that most of her students had experienced so much failure in mathematics that they felt defeated and completely uninterested in trying things at which they might fail again. "I'm all for keeping standards high, but I'm also for keeping it realistic," she said in her year 7 interview. "Students feel like it's too much. It's overwhelming. So if you can't build small increments of success then I don't care how high your standards are...they're gonna shut down." She also felt that CMP2's 'discovery' format was at an unnecessarily high level, and considered this particularly problematic with her regular classes. "No, they're not motivated," she said in her year 7 interview. "So if they're not motivated to begin with then they're not going to be motivated to discover anything."

In Lois's view, building "small increments of success" meant starting with simpler problems, providing students with multiple examples to help them identify problem types and appropriate solution methods, and motivating them through games, timed quizzes, and mnemonic devices. To accomplish this, Lois used CMP2 as a supplementary resource, and drew

heavily on other materials from other textbooks – as she imagined the CMP2 publishers had intended.

According to Lois, a major source of her students' lack of motivation was the prevalent focus on year-end test scores that were constantly “crammed down [students'] throat, like ‘look, see you didn’t’ pass, you didn’t pass, you didn’t pass....” (year 6 interview). Given her students' histories of failure, and what she perceived was a strong school-wide emphasis on analyzing student test data, Lois felt an enormous obligation to ensure that her students to score well on district benchmark assessments and year-end tests. By year 5, Lois had developed an effective routine for improving her students' outcomes on these assessments:

I look through the questions that are tricky and why they are tricky for the students, and then...make my own like worksheet or review with a bunch of those kind of things on it. So it's less about the [standard] and more about like how can I make a bunch of other questions that look like that, so the next time they see one like that, it won't be tricky....my improvement's been the greatest in my grade level, so I know my principal is happy with me right now (year 5 interview).

Lois also felt that it was important to give students structured routines for decoding the kinds of word problems they were likely to encounter on the state test. She used mnemonic devices to teach students to translate word problems into formulas by circling and crossing out ‘clue’ words that would indicate which procedure was called for. She taught these strategies routinely to her students – in effect, turning scenario-based word problems into simplified tasks that students could complete procedurally. In year 7, Lois found an additional instructional routine to address the motivation problem that she had noted in previous years: “Quick Quizzes.” Three times a week, she would time students as they worked on Quick Quizzes of multiplication

facts or procedural drills, and would mark their papers with the exact time it took them to finish the entire sheet. She admitted that her use of quick quizzes contradicted her own kind of vision of high quality instruction, but justified its place in instruction. “If you look at [the quick quizzes] you wouldn’t say ‘oh they’re higher level, blah blah’...the fact that they can get excited to participate has to start somewhere...from that I can build to where they can really, you know, [do] higher level thinking and all that...but you gotta start with that” (Lois, year 7 interview). Lois’s IQA data indicates though she may have wanted to eventually address “higher level” math with her students, her focus on recall and procedures was consistent each spring.

***Lois: supports for her learning.*** Similar to Mario, Lois was often frustrated by district professional development and school-based TCT. However, whereas Mario felt that the sessions were unhelpful because many of the attendees were not interested in learning more about instruction, Lois was frustrated by what she saw as time wasted talking about practice that could be better spent planning and organized for upcoming classes. For example, in year 7 she described attending a PD session in which district math leaders had teachers compare the new, more rigorous state standards with the former standards for their grade level. “That was awful,” she lamented. “Why are you even spending a whole day telling me what’s new?...I would rather go ahead and take what’s new and let’s say ‘what does that look like’ and give us a calendar for September and say, okay...figure out what your calendar might look like with the new [standards]” (year 7 interview). Similarly, she was outspokenly frustrated by her school’s adoption of the lesson planning protocol for use in TCT, which required an entire planning period to be spent thinking about only one day’s lesson rather than getting an entire week paced out and planned with time to run copies. After complying with the requirement to use the protocol for the first few meetings of the school year, she and the other 7<sup>th</sup> grade teachers finally

asked the department chair to be excused from using it, and were allowed to return to their usual form of planning. “We’ll get a little notecard,” she said, describing meetings that she found productive, “and we’ll put like the key topics for the week and where it is in the book, the CMP2, what lessons, and we decide ‘do we want to go with that or not? Do we want to compress, you know, 4.1 and 4.2?’ And then...we’ll decide who has what resources. And then...running our copies. And so that’s been good.”

Lois did not work closely with an instructional coach during years 5-7. Though the part-time math coach, Harry, did observe and co-teach in her classroom a few times in year 5, Lois reported feeling somewhat awkward in interactions with him during/after these observations. She did not have very much respect for him because of for his lack of math certification and lack of experience compared to her own. “How can you be the content specialist if you haven’t had that much classroom experience yourself, you know?” (year 5 interview). Unlike Mario, Lois never worked closely with Malcolm, Keisha, or Vera during years 5-7, though she did report that the school’s algebra and data coach, Trish, came into her classroom frequently during year 6 because, in Lois’ view, she was more “approachable” than other teachers in the department (year 6 interview). Lois appreciated her suggestions and contributions to the class – perhaps because Trish was also a very procedurally-oriented math teacher.

By year 7, when District B was reviewing new curricula to adopt in place of CMP2, Lois was relieved. “People in my team in 7th grade have said, you know, that we’re looking forward to the CMP2 going. It’s a good curriculum if it’s in conjunction with other curriculum or for [students enrolled in] honors, but for our kids we spend a lot of time pulling other resources” (year 7 interview).

**Summary.** In many ways, the contrasting cases of Mario and Lois at Two Lagoons parallel the contrast between Diane Cathy and Roger at Cypress. Like Diane and Cathy, Mario recognized that his goals for his students were different from those of his colleagues – that he was more interested in maintaining ambitious goals for their learning than lowering his goals to make CMP2 more “practical.” Though he did not necessarily identify with CMP2 whole-cloth as Diane and Cathy did, his problems with the curriculum were primarily about providing accessible problem contexts to his students, whose lives he did not see represented in the CMP2 books as they were written. This was not an insurmountable problem, and faced with no better option, Mario found ways to adapt the contextual features of CMP2 tasks for his students in ways that maintained the conceptual richness of the tasks while also capturing his students’ interest.

By contrast, Lois believed that CMP’s ambitious goals for student learning were unrealistic for her students, and identified motivating her students as her primary challenge. Both Roger and Lois were more often frustrated by available supports than they were motivated to engage with them, because they had very different goals for their own practice. While Lois did not disagree with her district’s vision of instruction as strongly as Roger did, she did qualify that vision of instruction as appropriate only for *certain* students – those who had mastered their basic skills, who were motivated to succeed in mathematics, and who were likely enrolled in honors classes. Thus, although Lois agreed in principle with the organization of CMP2, she (like Roger) felt it was unnecessarily high level for her students. Lois felt her students needed procedurally-oriented instruction that would enable them to experience success, by giving them “simple problems before more difficult ones” (year 7 interview). Similar to Diane and Cathy, when Mario encountered difficulties implementing CMP2, he was motivated to seek out advice

from others with expertise and experienced using the curriculum effectively. By contrast, Lois, who had set aside her ambitious vision in light of her perception of lower-level students' needs for basic skills practice, dreaded any PD sessions or TCT activities that focused on CMP2 investigations. In Lois's view, the district's goals for instructional improvement were irrelevant in light of her students' capabilities; time would be better spent planning what to cover on which days, and getting copies run.

**Anna (improver) and Roxanne (stable) at Park Falls.** Whereas the prior two cases illustrate differences in teachers' instructional improvement that seem more related to their individual expertise, identification with reform efforts, and engagement with supports, Anna and Roxanne had access to quite different levels of support for instructional improvement. Anna, who became a math coach in her second year as a math teacher, had much greater access to others with instructional expertise, and likely had higher quality supports for instructional improvement than did Roxanne. Anna's role as a coach included regular PD provided by district math specialists with significantly more expertise. Roxanne, meanwhile, had very little support beyond infrequent district PD, and over both years, maintained fairly unproductive views of how to support her 'regular' students who struggled with rigorous tasks.

***Anna: vision and identification.*** Anna started teaching math at Park Falls in year 2; she had previously taught middle school social studies for four years at another middle school. During her first year as a math teacher, Anna found CMP2 to be quite different from the instruction she had experienced as a student, and challenging to implement given the school's 45-minute periods. In year 3, when the math coach at Park Falls left the school, Anna was asked

by the principal to take her place since she had the most years of teaching experience in the department.

Anna had a relatively traditional VHQMI in year 2, which improved by year 4 as did her classroom instruction. Over those same years, however, Anna's VSMC – in particular, her descriptions of her approaches for addressing student learning difficulties – remained fairly unproductive. She frequently described the numerous structured tutoring programs she coordinated, in which students were pulled out of electives or other courses based on the specific test objectives they had scored poorly on or specific subgroups of which they were members.

Unlike Diane, Cathy, and Mario, who were (1) very explicit about their own commitments to using rigorous, conceptually-oriented instruction with their students, and (2) very aware that this made their classrooms and instructional choices look different than those of their colleagues, Anna did not clearly identify with the district's vision of instruction or the CMP2 materials specifically. Anna was generally less forthcoming than the other teachers in this study about her own views on school or district policies, including CMP2 and the forms of instruction it promoted. Unfortunately, coaches were not asked some questions which elicited the clearest indications of teachers' personal teaching identification ("In your view, is CMP appropriate for all students?"). Thus it is unclear whether Anna's steady instructional improvement was accompanied by personal motivation to see her students succeed with CMP2 and ambitious goals in general, or whether she was simply 'going along' with CMP in order to fulfill the district's expectations of her as a coach and district representative at Park Falls.

***Anna: supports for improvement.*** In her first year teaching math (year 2), Anna turned regularly to her math coach, who also taught 6<sup>th</sup> grade math, for support and advice. However, prior MIST analyses of this coach's expertise indicate that she was not particularly accomplished

in inquiry-oriented instruction, and her VHQMI was not much more sophisticated than the other teachers in the school. Moreover, Anna's work with her coach seemed to focus primarily on helping Anna develop routines for better classroom management, such as points systems and a popsicle sticks system designed to encourage student participation.

During years 2-4, Keisha, a district math specialist to whom Anna and a number of other math coaches reported visited Anna's school occasionally. Anna named Keisha as her chief source of advice and support during years 2-4, both in matters of teaching and coaching. In year 2, Anna described Keisha's role as a reporter of good ideas – e.g., telling teachers about strategies that were working well at other schools, such as having students create personal math glossaries to keep at their desks. In year 3, Anna described getting ideas from Keisha for resources to use with struggling students. Despite Keisha's significant expertise in math instruction, however, it's not clear that Anna ever received one-on-one coaching support from Keisha in her own classroom that was similar to the coaching that Diane, Cathy, or Mario received.

TCT at Park Falls during years 2-4 consisted of department meetings that were focused primarily on low-level planning/pacing/data analysis activities. These TCT meetings were facilitated by the math department chair, who had previously taught CMP2 at another school for several years and had become relatively accomplished at selecting and implementing high-level math tasks in her own classrooms in ambitious ways. However, despite the facilitator's expertise, Anna's and the teachers' descriptions these meetings indicate that they did not focus on instructional practice. These meetings were instead spent “examining data to identify students for tutoring, which [standards] need[ed] to be targeted, and what warm-up problems should be selected for the upcoming lessons” (Park Falls school summary form, year 3). The Park Falls

teachers also spent a great deal of time analyzing district and state assessment data to determine which standards to target when tutoring African American students, the sub-group that the school focused on the most. Because teachers from all three grade levels attended math department meetings, the department chair herself observed that “it’s a little difficult to get any kind of real gain out of [looking at data in TCT]... ’cause everyone’s’ got different data” (Park Falls department chair, year 3 interview).

Unlike the administrators at Mario’s and Diane and Cathy’s schools, the administrators at Park Falls did not seem to have any particular instructional expectations for math teachers beyond raising student achievement. None of the Park Falls teachers in our sample reported that the principal had any strong expectations for any particular teaching strategies or for the materials they should use during math instruction, and in Anna’s first and second year at Park Falls, both she and other teachers said they were very rarely observed. Though the principal set aside time during school-based Stipend Days to have district personnel lead PD sessions, they were usually content-general sessions (e.g., on manipulatives, accountable talk, flipcharts, grouping strategies, etc.) attended by the entire faculty.

Thus, TCT, coaching, and administrator expectations did not appear to have supported or pressed for steady improvements in Anna’s instruction over time. However, as a coach, Anna did participate in (and even facilitate) a significant amount of district math PD that may have been a significant support for her learning. Anna reported attending at least two different CMP institutes during years 2-4, which included time to work through a number of 6<sup>th</sup> and 8<sup>th</sup> grade CMP2 investigations. Beginning in year 3, when Anna took the position of coach, she also began attending monthly coach meetings that lasted for a half-day each. She said that these meetings involved (1) looking at, proofing, and improving the district-based math assessments, (2)

receiving materials and PowerPoint slides for upcoming district PD sessions that coaches would facilitate in pairs, and (3) discussing books they were reading about improving their coaching practices. District math leaders' descriptions of these meetings indicate that they also included role-playing activities in which the coaches would take turns rehearsing how to talk to a teacher after observing instruction. According to district math leaders, these monthly meetings for math coaches were intended to support the coaches to become more than just district liaisons, and to develop coaching practices that they were to take back to their schools. For example, as a coach, Anna was expected to choose one (in year 3) or two (in year 4) teachers to work with one-on-one during the year by enacting a coaching cycle, in which she co-planned a lesson with the teacher, observed the teacher's instruction, and debriefed with them afterward. Both years, Anna described enacting these cycles with new teachers at her school, who she felt needed support to feel welcome in the school and develop routines for effective classroom management.

Anna did not describe ever doing mathematics in these monthly coach meetings, nor did she describe talking about her own classroom instruction at these sessions. Nonetheless, it is possible that these meetings provided Anna with opportunities to identify ways to improve her own instruction. For example, the sessions spent developing coaching practices to support teachers' in improving their own instruction may have supported her development of a more ambitious vision for her own practice.

**Roxanne.** Roxanne also began teaching middle school math at Park Falls in year 2; she had previously been a science teacher at a different school for several years. In the two years she taught at Park Falls, Roxanne primarily taught advanced courses in Algebra and Geometry meant to mirror the content taught at the high school level. Surprisingly, Roxanne's MKT was below

average among teachers in the district – even in Algebra, which she spent much of her day teaching.

Both years she taught at Park Falls, Roxanne only taught one “regular” 8<sup>th</sup> grade math class using CMP2. Both years, she named this class as her biggest challenge, and often remarked on how different those students were from the advanced students she worked with the rest of the day. She described how her ‘regular’ students had behavior problems, were less capable of working in groups, and struggled to do the kinds of reasoning, explaining, and justifying that CMP called for. She also said that these students were not accustomed or able to do as much writing and talking about math as CMP asked for. Her way of addressing these problems was to lower the cognitive demand of tasks or activities rather than providing supports that would enable students to engage meaningfully in rigorous tasks.

Though Roxanne did not voice as much opposition to CMP2 as either Roger or Lois, it seems that the challenges she encountered in implementing CMP2 with her regular class proved overwhelming. By the end of year 3, Roxanne indicated on the MIST teacher survey that the primary math curriculum at her school (CMP2) was “not at all consistent” with her personal beliefs about effective teaching methods.

Though Roxanne never articulated the specific problems she saw with CMP2 in interviews, descriptions of Roxanne’s instruction by the math department head suggest that Roxanne’s vision of instruction may have been influenced by the courses she was teaching. Stephanie, the math department head, remarked in her year 2 interview that most of the math teachers in the building shared a fairly ambitious vision of instruction – except for Roxanne. She said that Roxanne’s “vision is much different because she’s teaching an upper level course and

traditionally, high school courses...they just haven't followed suit...and I understand that, so she teaches in a more traditional way" (Park Falls department head, year 2 interview).

This description by Roxanne's department head matches prior MIST analyses of Roxanne's instruction: her IQA data from years 2-3 indicates that her typical approach was to use fairly routine, procedural tasks with her students, demonstrate steps to solve them, and monitor students when they subsequently they worked on practice problems. Roxanne's procedurally-oriented instruction in her regular class, particularly in her second year, may have also been a consequence of a narrow focus on improving students' test scores. Roxanne explained in her year 3 interview that her principal was concerned that many her advanced students had unexpectedly failed the state test the prior year. This intense pressure she felt to improve the test outcomes of her advanced classes, coupled with her distaste for CMP2, may account for her use of lower-level, procedural tasks in her regular class during both years she taught math at Park Falls.

***Roxanne: supports for learning.*** Roxanne, like Anna, had very few ongoing opportunities to work closely with a coach with more expertise than her. During years 2 and 3, the only other teachers with instructional expertise at Park Falls were the Stephanie (the department chair), who taught a different grade level, and Anna, whose vision and practice were only slight more sophisticated than Roxanne's at the time. In their roles as TCT leader and coach (respectively), neither seemed to coordinate or facilitate high-leverage activities with Roxanne aimed at supporting her development of more ambitious goals or practice. Roxanne did describe being observed at her own request by the coach during her CMP2 class in year 2 (the coach before Anna). However, these observations did not seem to focus on helping her problematize her current forms of practice; rather, she said the coach was focused on "objectives, we've got

our word wall up to date, that we are following along with...the curriculum, student work is posted..." (Roxanne year 2 interview). The coach did give her feedback following these observations, but it focused on how to engage particular students who were not paying attention during her lessons. More generally, she said that her coach's usual feedback was "I don't ever have to worry about you; you're always doing what you're supposed to do" (Roxanne, year 2 interview). In year 3, Anna also observed Roxanne teach a CMP2 lesson; however, Roxanne did not describe any specific feedback she was given after those observations. In year 3, Roxanne said she often asked Anna for advice or insight because she had taught all three grade levels of CMP (and they were both teaching 8<sup>th</sup> grade that year). However, beyond these informal conversations between classes, Anna did not regularly provide Roxanne with coaching support in the form of modeling, co-teaching, or planned cycles of planning, observation, and feedback.

As described above in Anna's case, TCT at Park Falls was generally involved low-level pacing and planning activities, and analyzing student assessment data in order to group students for the many forms of tutoring that were organized before, during, and after school. Roxanne did not seem to view these meetings as very helpful. Even when they did have time to discuss 8<sup>th</sup> grade students or curriculum in particular, she said that did not get much out of the discussions since she only taught one of the 'regular' 8<sup>th</sup> grade classes, and she was the only teacher who taught the advanced classes.

Given the lack of ongoing support Roxanne received to improve her practice, her lack of experience or expertise in math instruction, the school's emphasis on tutoring solutions to differential student testing outcomes, and the pressure she felt from her principal to improve her advanced students' scores, it is understandable that Roxanne did not develop more ambitious practice.

*Summary.* In contrast to the cases at Cypress and Two Lagoons, the differences between Anna’s and Roxanne’s vision and identification are less pronounced. Neither explicitly aligned themselves with the goals and practices they saw represented in the CMP2 curriculum or the district PD. Furthermore, both teachers described fairly low-level supports for struggling students, such as race-based class assignments and tutoring. Roxanne in particular equated students’ enrollment in regular and advanced classes with their capability to participate in group work, to have mathematical discussions, and to work with less explicit guidance from their teacher.

The most significant difference between Anna and Roxanne appeared to be the amount of support they received to improve their instruction. Anna, as the math coach, had ongoing support from Keisha and participated in monthly professional development sessions led by district math specialists with significant expertise in instruction and CMP2. As part of this PD, she read two books on math coaching, had opportunities to rehearse coaching conversations with teachers, and used math-specific observation forms with teachers at her school. Given the expertise of the math specialists leading these sessions and the nature of some of the activities in the sessions, it is likely that they were instrumental in supporting Anna’s development of a sophisticated vision and more ambitious instruction during years 2-4. Roxanne, meanwhile, had much less support for improving her instruction. While she found the summer CMP2 institutes encouraging and well-facilitated, she was also discouraged to find that the rosy picture of CMP2 painted at these institutes was much more difficult to accomplish with her own students. And with no school-based expert available to model ambitious instruction with her students, or problematize her traditional, “high-school” style of teaching, Roxanne’s instruction remained procedurally-oriented in both years.

## Synthesis

The three comparative cases presented in this section illustrate the importance of both high-quality supports for teacher learning, and teachers' individual sense of identification with the vision of instruction advocated by those supports.

The contrasting cases of Anna and Roxanne at Park Falls illustrate the importance of providing ongoing support for teachers to develop ambitious instruction. In Roxanne's case, although the summer PD sessions she attended appeared to be high quality, they were not sufficient for supporting her to develop ambitious goals for all of her students or instructional practices that would help students reach those goals. Furthermore, Roxanne's survey responses regarding CMP2 indicate that she did not come to identify with the teaching methods represented in the curriculum – perhaps because the bulk of her day was spent teaching upper-level courses characterized by traditional, procedurally-oriented teaching methods. Taken with my finding that both improvers and decliners attended relatively high-quality PD, Roxanne's case further illustrates that teachers need more support than occasional pull-out PD if they are develop more ambitious practices – particularly as they grapple with challenging, reform-oriented curricula that they feel are beyond their own students' capabilities.

My analysis of the contrasting cases at Park Falls and Two Lagoons illustrates how a teacher's vision of instruction and views of students' mathematical capabilities likely contributed to their sense of identification with their district's reform efforts. Roger, who was outspokenly opposed to the district's implementation of CMP2, objected to the forms of instruction that he associated with the curriculum – in part because of the way it was implemented, and in part because he subscribed to a very traditional vision of instruction. His VHQMI both years indicated that he considered procedural demonstration and independent practice to be key aspects

of high quality instruction – and that ‘regular’ students in particular were not capable of anything more rigorous. Though his principal made a number of efforts to support the math department, and Roger in particular, to successfully implement the curriculum, Roger rebuffed or remained very wary of these supports because he saw them as attempts to force him to do things he did not believe in.

Lois, by contrast, described a fairly inquiry-oriented vision of instruction. Over years 5-7, Lois’s VHQMI was relatively high when compared with other teachers in the district; in her view, high quality lessons involved students solving challenging problems that required them to think, reason, and discuss – not just repeat procedures. However, Lois’s relatively high VQHMI was undermined by her unproductive VSMC: this high quality instruction was unrealistic for her students, who believed were unmotivated due to prior failure and lack of basic skills. Her instruction was consistent with the lower-level goals she set for students. Supports such as PD or TCT that focused on more ambitious goals or practices were frustrating for her as she wanted uninterrupted time for the real work of planning: selecting worksheets, pacing out the upcoming week’s standards, and running copies.

Lois’s and Roger’s cases both illustrate how teachers’ sense of identification with particular forms of practice heavily depends on their vision of high quality instruction *and* their views of what mathematics their own students are capable of doing with appropriate support. Lois’s case in particular illustrates how even teachers who have a relatively ambitious VHQMI may continue to see rigorous practice as beyond the capabilities of their own students, and thus draw on low-level tasks and activities for daily instruction rather than the kinds of tasks they themselves view as higher quality. Three of the decliners I analyzed for this study – Penny, Alishia, and Karyn – had a similar combination of high-VHQMI and low-VSMC. The fourth, Li,

had a fairly procedurally-oriented VHQMI *and* unproductive VSMC all three years of his decline. Taken together, these cases indicate that a teacher’s personal teaching identification – e.g., the forms of practice they hold themselves accountable to – is underpinned by her vision of instruction and her views of her own students’ capabilities.

The cases of Lois and Roger also illustrate how teachers’ own sense of identification with (or resistance to) their district’s goals for instruction may shape their engagement with available supports. Both Lois and Roger, like the improvers in their schools, had access to high-quality supports and other teachers or coaches with significant expertise in inquiry-oriented instruction. However, their own views on what *their* students needed and the best way to meet those needs as a teacher led them to engage with those supports very differently than their improver counterparts.

## CHAPTER V:

### Conclusions & Implications

#### Critical aspects of teachers' current practice

I initially conjectured that four primary indicators of instructional expertise would be related to teachers' development or decline in instructional quality: their prior forms of instruction, their math knowledge for teaching, their vision of high quality math instruction, and their framing of student learning difficulties.

It proved difficult to assess the influence of prior instruction adequately because the MIST dataset is limited to the years during and after which the participating districts first introduced CMP2 in middle schools. Thus, though several of the teachers selected for case studies were math teachers prior to MIST year 1, their prior forms of instruction are unknown. Furthermore, several of the teachers in this study were new to the profession or to mathematics teaching when they began as participants in the MIST project. Thus, it was not possible to analyze adequately how the case study teachers' prior instruction influenced the extent to which they engaged with and learned from supports like PD, coaching, and TCT.

While I did find variation between teachers in their MKT, this variation did not correspond systematically to changes in the quality of their instruction. In fact, three of the five teachers with significantly stronger MKT were decliners, and also had unproductive views of their students' capabilities in math. Meanwhile, the MKT of three of the four improvers remained fairly stable over the years their instruction was improving. This finding adds to existing research that though having advanced math knowledge for teaching may be important for teachers who

use cognitively demanding tasks during instruction, strong MKT does not necessarily make teachers more likely to do so (Wilhelm, 2014). For example, stable teacher Lois's MKT improved dramatically in years 5-7 when she was enrolled in a career advancement master's program. However, she maintained the view that the best way to support students who she viewed as unmotivated because of prior histories of failure was to build their confidence by giving them low-level, procedurally-oriented tasks and speed drills. On such tasks, there is little mathematical uncertainty to navigate with advanced MKT (e.g., little variety in student reasoning methods and few opportunities for novel problem solving approaches). Thus, though the classes she attended as part of her Master's program may have supported the development of her capacity to reason through mathematically diverse strategies, such diversity was unlikely to occur in her classroom because of her views of her students' needs, and how to meet those needs. Similarly, decliner Li's high MKT was likely useful his first year when he primarily used CMP2 tasks as the basis for instruction. However, what he saw as a lack of self-discipline and scholarly drive in his students led him to lower his expectations for his students, and the rigor of his instruction. These cases add further evidence that supports specifically designed to develop teachers' MKT may be more appropriate for teachers who have already come to identify with a more ambitious vision of instruction and recognize their own need for deeper content knowledge (Hill & Ball, 2004).

Looking across all the cases I examined, I also found that an ambitious vision of high quality instruction is necessary, but not sufficient, for motivating a teacher to improve their practice. Three decliners (Penny, Alishia, and Karyn) viewed inquiry-oriented mathematics, problem-solving scenarios, and conceptual discourse as features of high-quality mathematics instruction, as did Lois (stable). However, none of them viewed this vision of instruction as

attainable for all of their students, and they therefore reduced the cognitive demand of typical math tasks for their students.

When comparing improvers with decliners, and with their stable counterparts, the key difference that I found was that improvers had both an ambitious *vision* of instruction, and viewed that these ambitious goals were appropriate for even students who were struggling. This combination of an ambitious vision and productive views of student capabilities seemed to underpin these teachers' personal identification with ambitious instruction, and the accountability-to-self they felt when they could not attain their ambitious goals with all their students all the time. For example, Mario (improver) described this accountability in terms of a "burden" to serve struggling students while maintaining ambitious goals for their development of conceptual understanding. The improving teachers' ambitious visions of instruction, and their views that such a vision was appropriate for all *their* students seems to explain why they saw the inquiry-oriented curriculum their districts had chosen (CMP2) as an appropriate, if imperfect, tool for accomplishing this vision in their own classrooms. Mario, for example, remained critical of CMP2 across years 4-7, primarily because task scenarios and contexts were removed from the lives of his primarily Latino students who had been raised in urban environments. However, he found the mathematical rigor of CMP2 and its emphasis on mathematical 'processing' made it the best available resource for reaching his goals with students, despite the time required to modify scenarios and contexts for his students.

### **Critical Supports for Teacher Learning**

Based on prior research in teacher professional development, instructional coaching, and TCT, I conjectured that case teachers' access to high quality supports would account, in part, for

their changes in practice over time (e.g., that improvers likely had greater access to high quality supports than decliners). Additionally, I conjectured that the expectations of teachers' school leaders would also influence whether they were likely to improve their instruction over time. For example, that decliners may have perceived that their principals wanted to see more focus on low-level skills building in alignment with upcoming state tests.

In comparing improvers with decliners, I did not find any pronounced differences between groups in the quality of the professional development they attended over time, nor in the kinds of activities they typically did in TCT. I also found no clear differences between improvers and decliners in their perception of their school leaders' instructional expectations.

Nearly all the teachers I studied – improvers, decliners, and stable procedurally-oriented teachers – reported attending multiple pull-out PD sessions that focused on improving their vision and practice related to implementing CMP2. Though the case teachers gave mixed reviews of the quality and helpfulness of these sessions, even these reviews did not differ according to case type; that is, improvers were no more likely to report that the sessions were helpful, and decliners and stable teachers were no more likely to report that they were unhelpful. This finding suggests that high quality pull-out PD may be necessary, but insufficient, for supporting teachers' development of ambitious practice over time. Working through math tasks together and planning lessons or units with other teachers may have been an important support for improvers' development of practice. However, it seems that engagement in these PD activities did not impact decliners' and stable teachers' unproductive views that their own students were not capable of engaging in inquiry instruction. While it is possible that PD targeting these specific issues may have had different outcomes, I did not have opportunity to investigate any such cases.

There was also no systematic variation in the quality of teachers' opportunities to learn in their regular TCT meetings. While there were some reports of potentially high-depth TCT activities that went beyond sharing "tips and tricks," these reports were exceptions. More commonly, teachers reported planning for the upcoming instruction during TCT by coordinating their pacing schedules with other teachers in their grade level. At many schools, TCT was also used to look at school-wide trends in student assessment data in order to identify students and topics for additional test-related tutoring. While improvers *were* more likely to have a TCT facilitator with relatively more instructional expertise than the teachers in the building (usually the math coach), these facilitators were no more likely to focus this time on deeper discussions of instruction or student learning. This prevalence of low-level TCT activities is likely due to a lack of support or accountability for teachers to use the time for any specific purposes. During years 1-5 of MIST, neither district communicated any clear expectations (or provided any particular tools or supports) for what teachers should be doing during TCT. Though District B provided the lesson planning protocol for TCT in years 6-7, there was no coordinated support for TCT leaders to learn to use this protocol productively.

I also found no clear differences between improvers', decliners', and stable, procedurally-oriented teachers' perceptions of their principals' instructional expectations. Across cases, teachers reported that their principals were generally concerned with surface-level aspects of instruction, such as whether and where the lesson objective was posted, and how many students were actively participating in the lesson. Though several of the principals had a markedly ambitious vision of instruction, they either did not communicate such expectations to their teachers (e.g., the principals at Alishia and Penny's schools - decliners), or these expectations made little impact on teachers' practices. For example, while the principals at Two Lagoons (in

year 4) and Cypress did communicate clear expectations that teachers only use CMP2 for primary math instruction, both Lois and Roger (stable) found means to skirt these expectations – either by hiding their use of supplementary materials, or by walking their students through CMP2 investigations step-by-step as they would with traditional textbook tasks.

Where I did find important differences between improvers and decliners was in their access to (and work with) instructional coaches who had significant expertise in ambitious instruction (e.g., an ambitious VHQMI, productive VSMC, and in many cases, had developed ambitious instructional practices themselves before becoming a coach). Diane and Cathy’s (improvers) coach, Karon, worked with them and other teachers in the building (including Roger, stable) as a support for their implementation of CMP2. This support included co-planning lessons, modeling lessons for them with their own students, and providing them with feedback on their instruction. Similarly, Mario (improver) was mentored his first year by both Keisha (a district math specialist) and Vera, his AP who had significant expertise in ambitious math instruction and had previously worked as a district specialist. In later years, he was supported by Malcolm, who visited Two Lagoons regularly to answer teachers’ questions about CMP2 and provide them with feedback on their instruction. Anna (improver) also received support from Keisha during years 2-4 when she worked at Park Falls as a part-time math coach. Though her work with Keisha (and the monthly coaching meetings and PD sessions she attended) focused on her development as a coach, this support may have also been a critical factor in Anna’s development of ambitious instructional practice.

## **Implications for Future Research in Teacher Learning**

The findings of my analysis suggest several avenues for future research. First, this study calls for larger scale, follow-up studies. This study purposefully drew on outlying cases of instructional improvement and decline in order to generate conjectures about teacher learning that may generalize to other cases of teacher learning in their institutional settings (Flyvbjerg, 2006). Future research should therefore explore the generalizability of these findings by developing empirically grounded means of testing similar relationships in larger samples of teachers.

One avenue for future inquiry using the MIST dataset would be to test the conjecture that only teachers who have developed both an ambitious VHQMI and productive VSMC come to identify with ambitious goals for instructional improvement promoted in their school or district. Testing this conjecture in the MIST sample would involve first developing a methodologically reliable and valid measure of teachers' identification with or resistance to their district's reform efforts – for example, a qualitative coding scheme that could be used reliably to generate categorical codes for all teachers in the dataset. These codes could then be used in logistic regressions to test whether VHQMI+VSMC is a significant predictor of teachers' identification with ambitious goals for instructional improvement. Similarly, these codes could also be used to test the conjecture that teachers who identify with an ambitious vision are more likely to engage with available supports for developing ambitious practice (e.g., whether teachers who identify with an ambitious are more likely than their peers to attend voluntary PD or seek out the advice of a math coach with significant instructional expertise).

A second and related avenue for research using the MIST dataset would be to explore other conditions necessary for teachers to come to identify with an ambitious vision. While a

reform-oriented VHQMI and productive VSMC might to be necessary for identification, it is not clear that they are sufficient — that every teacher with strong VHQMI and VSMC will necessarily identify with ambitious goals for their own practice. For example, it is conceivable that a teacher who has an ambitious vision of instruction and sees her own students as capable of engaging in rigorous math does not come to identify with this vision because she knows that students will only be held accountable on state tests for mastering basic procedural skills. In this case, she may weigh the cost of spending additional preparation and class time on rigorous tasks against the lack of measurable impact on students' test scores, and ultimately decide that inquiry-oriented instruction is not worth the cost.

Future research should explore this and other possible reasons why VHQMI and VSMC may be necessary but not sufficient predictors of teachers' personal teaching identification. One way to approach this question in the MIST sample would be to pool teachers with both sophisticated VHQMI and productive VSMC, and then compare the school conditions and relevant instructional experiences of those who identify with an ambitious vision to those who do not. Such a comparative analysis could generate testable conjectures about additional factors that contribute to a teacher's professional identity – conjectures which could guide the development of future interventionist work.

### **Implications for Research and Practice in Professional Development**

My findings about the relationship between key aspects of teachers' knowledge, their identification with particular forms of practice, and the supports they receive to develop improved instruction also have implications for the design of teacher professional development – including pull-out PD, coaching, and TCT.

First, these findings suggest the importance of developing teachers' sense of personal identification with district reform efforts by supporting their development of a sophisticated vision of high quality instruction and productive views that their students' mathematical capabilities. While many practice-based approaches to teacher professional development may implicitly include these goals for teachers' VHQM and VSMC, my findings suggest that it may be prudent to attend to these aspects of teachers' expertise explicitly during the initial phase of instructional improvement efforts.

**Developing an ambitious vision.** Supporting teachers' development of a more sophisticated instructional vision of might entail, for example, engaging them as learners in problem-solving tasks that require them to explain or justify their thinking in groups and in larger discussions. Experiencing ambitious mathematics instruction from the point of view of the learner, followed by participating in reflective discussions of the role that tasks, discussion, and the teacher played in creating the experience, might be a rich opportunity for teachers to problematize their current forms of practice, and consider the affordances of more ambitious instruction (Jackson, Gibbons, & Dunlap, in press; Elliott et al., 2009; Franke & Kazemi, 2001; Koellner et al., 2007).

**Anticipating challenges and developing productive views of diverse students.** As one of the case study teachers, Roxanne, pointed out, things do not typically play out in the classroom as they do when teachers play the role of students at PD sessions. Teachers are bound to encounter challenges associated with students' lack of experience in discourse-rich, conceptually-oriented classrooms. As Roxanne described, "it is difficult to get [students] to write

and to get them to discuss the math problems because they are just not used to it” (year 2 interview). Furthermore, teachers are likely to encounter the frustration of students who have typically been successful in procedural activities who exclaim, as Lois reported, “I don’t want to discover it. Just tell me what it is....show me some examples, like three different ways [I] might see this problem, then I know how to recognize, ‘oh, that’s gonna be this kind of problem’” (year 5 interview). Even teachers like Cathy, who had used CMP2 for several years, found it challenging to implement CMP2 with students who were not use to persisting in problem-solving; for Cathy, this resulted in classroom management problems she had not encountered at her previous school. Teacher who encounter these and other challenges in their own classrooms might well conclude, as Karon, the coach at Cypress described, that their students are somehow different and that rigorous math instruction is not appropriate for them.

It is predictable that most teachers will encounter these and other challenges when attempting to implement ambitious tasks in their own classrooms. The findings I had reported suggest that teacher professional development (e.g., pull-out PD, coaching, and TCT) should include opportunities for teachers to problematize their unproductive views about whether their students are capable of engaging in rigorous instruction. For example, such opportunities might involve working directly with (or observing) their own struggling students, or students they view as having similar difficulties in mathematics, engaging in the rigorous instruction.

Nelson, Lesseig, Slavit, Kennedy, and Seidel (2015) describe two one-week summer PD sessions conducted with high school science and math teachers provides a useful model in this regard. In the first week, teachers engaged in STEM “Design Challenges” written by the research team for high school students. In these design challenges, teachers worked in groups with other teachers in cycles of engineering design, rapid prototyping, and design improvement. Afterward,

they discussed the affordances of engaging students in similar work, and the implications for teaching practices that would facilitate their students' productive engagement. The participating teachers then identified the students who they had found it most challenging to engage in rigorous tasks during the prior year. The PD facilitators then invited those students to participate in similar STEM design challenges during a second week of PD later in the summer. During that second week, PD facilitators divided the students into groups, placing each participating teacher in a group to work as an equal with their students (e.g., they were not to instruct or guide students beyond working as another student might in the group). Slavit and Nelson report that many of the teachers admitted to being shocked at their students' ingenuity, persistence, and conceptual understanding – many of whom had learning disabilities and language barriers that typically led to difficulties in their regular classwork.

**Enacting initial, high-leverage instructional routines.** Once teachers have developed an ambitious vision of instruction *and* have come to see their students as capable of engaging in rigorous mathematics, they will need support to then *enact* more ambitious practices in their own classrooms. These supports might include the introduction of some initial routines or instructional activities that are easy to adapt for a range of mathematical concepts, but that provide significant opportunities for students to engage in mathematical discourse and develop conceptual understanding. For example, high leverage routines in elementary or middle grades might include brief (10-15 minute) discussions of strings of related problems (Parrish, 2010), or choral counting activities that build on students' understanding of number relationships to build fluency (Kazemi and Hintz, 2015; TEDD.org, 2016).

In sum, this proposed sequence of teacher learning opportunities would be aimed at supporting teachers to develop an ambitious vision of instruction and productive views of their students' capabilities *before* supporting the development their instructional practices. This design principle builds from my finding that teachers' development of both sophisticated VHQMI and productive VSMC potentially underpins their identification with an ambitious vision of instruction. Future research is needed to test these design principles and elaborate upon them with productive designs for PD (broadly defined) that support teachers' development in these areas. Given that there is little research that has explored means of improving teachers' VSMC (Jackson, Gibbons, & Dunlap, in press), design research methodology may be appropriate (Cobb, Jackson, & Dunlap, 2014). For example, the first step in this line of work might involve a PD design experiment with a small number of teachers aimed at supporting them to come to identify with ambitious goals for their own instruction (e.g., Gresalfi & Cobb, 2011). Such an experiment might begin by adapting the PD designs used by Slavit and Nelson (2016) to support teachers to develop and maintain ambitious goals for their most challenging students.

Further research is also needed that specifically attends to the role that TCT and coaching can play as part of this broader agenda for supporting teacher learning. For example, additional research is needed that investigates how coaches can support teachers to recognize and problematize the unproductive views they may have of students who have had the most difficulties engaging in rigorous mathematics tasks – perhaps through modeling, or through targeted feedback following classroom observations. Pursuing this agenda in TCT might involve investigating how to support teachers' learning during relatively common TCT activities such as planning upcoming math lessons together or analyzing student assessment results. It might be feasible, for instance, to support relatively accomplished TCT facilitators to focus the group on

understanding *why* students may struggle on particular tasks and items, and considering what can be done to get them back ‘in the game’ of ambitious math instruction (Brasel et al, 2016).

The findings of my analysis indicate that over the course of just a few years, teachers can become quite accomplished in selecting and implementing rigorous tasks and activities in diverse classrooms. Thus, while supporting district- or school-wide instructional improvement would require substantial investment in supports for teacher learning, it is not an impossible task.

## REFERENCES

- Ai, X. & Rivera, N. (2003). Linking ideas to practice: Effectiveness of coaching upon teacher practice. In *Annual Meeting of the American Educational Research Association, Chicago, IL* (pp. 1–46).
- Ball, D. L. (1996). Teacher learning and the mathematics reform: What we think we know and what we need to learn. *Phi Delta Kappan*, 77(7), 500.
- Ball, D. L. & Bass, H. (2000). Interweaving content and pedagogy in teaching and learning to teach: Knowing and using mathematics. *Multiple Perspectives on the Teaching and Learning of Mathematics*, 83–104.
- Ball, D. L. & Cohen, D. (1999). Toward a practice-based theory of professional education. *Teaching as the Learning Profession San Francisco: Jossey-Bass*.
- Ball, D. L., Hill, H. C. & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide?
- Booher-Jennings, J. (2005). Below the bubble: “Educational triage” and the Texas Accountability System. *American Educational Research Journal*, 42(2), 231–268.
- Brasel, J., Garner, B., Kane, B., & Horn, I. (2015). Getting to the Why and How. *Educational Leadership*, 73(3).
- Bryk, A. S. (2010). Organizing schools for improvement. *Phi Delta Kappan*, 91(7), 23–30.
- Campbell, P. F. & Malkus, N. N. (2009). School improvement through elementary mathematics coaches: Impact on teacher beliefs and student achievement. annual meeting of the American Educational Research Association, San Diego, CA.

- Campbell, S. S. (2012). *Taking Them Into the Field: Mathematics Teacher Candidate Learning About Equity-Oriented Teaching Practices in a Mediated Field Experience*.
- Cobb, P., Jackson, K. & Dunlap, C. (2014). Design research: An analysis and critique. *Handbook of International Research in Mathematics Education*.
- Cobb, P., McClain, K., de Silva Lamberg, T. & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher*, 32(6), 13–24.
- Cobb, P., Zhao, Q. & Dean, C. (2009). Conducting design experiments to support teachers' learning: A reflection from the field. *The Journal of the Learning Sciences*, 18(2), 165–199.
- Cohen, D. (2011). *Teaching and its Predicaments*. Cambridge, MA.: Harvard University Press.
- Crockett, M. D. (2002). Inquiry as professional development: Creating dilemmas through teachers' work. *Teaching and Teacher Education*, 18(5), 609–624.
- Dunlap, C., Hunter, S., Rosenquist, R. & Smith, T. (Revise and Resubmit, July 2016). *Continuing the Discussion: comparing self-reports to classroom observations for evaluating PD effectiveness*.
- Elliott, R., Kazemi, E., Lesseig, K., Mumme, J., Carroll, C. & Kelley-Petersen, M. (2009). Conceptualizing the work of leading mathematical tasks in professional development. *Journal of Teacher Education*, 60(4), 364–379.
- Ermeling, B. A. (2010). Tracing the effects of teacher inquiry on classroom practice. *Teaching and Teacher Education*, 26(3), 377–388.
- Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *The Teachers College Record*, 103(6), 1013–1055.

- Franke, M. L. & Kazemi, E. (2001). Learning to teach mathematics: Focus on student thinking. *Theory into Practice*, 40(2), 102–109.
- Gainsburg, J. (2012). Why new mathematics teachers do or don't use practices emphasized in their credential program. *Journal of Mathematics Teacher Education*, 15(5), 359–379.
- Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F. & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915.
- Gibbons, L. K. (2012). *Examining Mathematics Coaches in Supporting Teachers to Develop Ambitious Instructional Practices: A Three-paper Dissertation Investigating Context, Knowledge, and Practice*. Doctoral Dissertation: Vanderbilt University.
- Glaser, B., & Strauss, A. (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine.
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E. & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *The Teachers College Record*, 111(9), 2055–2100.
- Grossman, P., Wineburg, S. & Woolworth, S. (2001). Toward a theory of teacher community. *The Teachers College Record*, 103(6), 942–1012.
- Hall, R. & Horn, I. (2012). Talk and conceptual change at work: Adequate representation and epistemic stance in a comparative analysis of statistical consulting and teacher workgroups. *Mind, Culture, and Activity*, 19(3), 240–258.
- Hiebert, J., Stigler, J. W., Jacobs, J. K., Givvin, K. B., Garnier, H., Smith, M., ... Gallimore, R. (2005). Mathematics teaching in the United States today (and tomorrow): Results from

- the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis*, 27(2), 111–132.
- Hill, H. C. & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, 330–351.
- Hill, H. C., Ball, D. L. & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 372–400.
- Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406.
- Hindin, A., Morocco, C. C., Mott, E. A. & Aguilar, C. M. (2007). More than just a group: Teacher collaboration and learning in the workplace. *Teachers and Teaching: Theory and Practice*, 13(4), 349–376.
- Horn, I. S. (2007). Fast kids, slow kids, lazy kids: Framing the mismatch problem in mathematics teachers' conversations. *The Journal of the Learning Sciences*, 16(1), 37–79.
- Horn, I. S., Kane, B. D. & Wilson, J. (2015). Making Sense of Student Performance Data Data Use Logics and Mathematics Teachers' Learning Opportunities. *American Educational Research Journal*, 0002831215573773.
- Horn, I. S., Nolen, S. B., Ward, C. & Campbell, S. S. (2008). Developing practices in multiple worlds: The role of identity in learning to teach. *Teacher Education Quarterly*, 35(3), 61–72.

- Howe, R. (1999). Knowing and teaching elementary mathematics. *Journal for Research in Mathematics Education*, 30(5), 579–589.
- Jackson, K., Cobb, P., Wilson, J., Webster, M., Dunlap, C. & Appelgate, M. (2014). Investigating the development of mathematics leaders' capacity to support teachers' learning on a large scale. *ZDM*, 1–12.
- Jackson, K., Gibbons, L. & Dunlap, C. (in press). Teachers' Views of Students' Mathematical Capabilities: Challenges and Possibilities for Ambitious Reform.
- Jenlink, P. M. & Kinnucan-Welsch, K. (2001). Case stories of facilitating professional development. *Teaching and Teacher Education*, 17(6), 705–724.
- Joyce, B. & Showers, B. (1982). The Coaching of Teaching. *Educational Leadership*, 40(1), 4–8.
- Kazemi, E. & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, 7(3), 203–235.
- Kazemi, E., Franke, M. & Lampert, M. (2009). Developing pedagogies in teacher education to support novice teachers' ability to enact ambitious instruction. In *Crossing Divides, Proceedings of the 32nd Annual Conference of The Mathematics Education Research Group of Australasia* (Vol. 1, pp. 11–29).
- Kazemi, E. & Hintz, A. (2014). *Intentional talk: How to structure and lead productive mathematical discussions*. Stenhouse Publishers.
- Kelley-Peterson, M. M. (2010). *Understanding Ambitious Mathematics Teaching Practice through Instructional Activities*. ERIC.
- Koellner, K., Jacobs, J., Borko, H., Schneider, C., Pittman, M. E., Eiteljorg, E., ... Frykholm, J. (2007). The problem-solving cycle: A model to support the development of teachers' professional knowledge. *Mathematical Thinking and Learning*, 9(3), 273–303.

- Lampert, M., Boerst, T. A. & Graziani, F. (2011). Organizational resources in the service of school-wide ambitious teaching practice. *Teachers College Record*, 113(7), 1361–1400.
- Lampert, M., Franke, M. L., Kazemi, E., Ghouseini, H., Turrou, A. C., Beasley, H., ... Crowe, K. (2013). Keeping It Complex Using Rehearsals to Support Novice Teacher Learning of Ambitious Teaching. *Journal of Teacher Education*.
- Leithwood, K. A. & Riehl, C. (2003). *What we know about successful school leadership: Executive Summary*. National College for School Leadership Nottingham.
- Levine, T. H. & Marcus, A. S. (2010). How the structure and focus of teachers' collaborative activities facilitate and constrain teacher learning. *Teaching and Teacher Education*, 26(3), 389–398.
- Lewis, R., Gibbons, L., Hintz, A. & Kazemi, E. (2015). Teacher Time-Out: A Way to Support the Collective Learning of Educators. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Lincoln, Y. S. & Guba, E. G. (1985). Establishing trustworthiness. *Naturalistic Inquiry*, 289–331.
- Little, J. W. (2002). Locating learning in teachers' communities of practice: opening up problems of analysis in records of everyday work. *Teaching and Teacher Education*, 18(8), 917–946.
- Little, J. W. (2012). Professional community and professional development in the learning-centered school. *Teacher Learning That Matters: International Perspectives*, 22–43.
- Matsumura, L. C., Sartoris, M., Bickel, D. D. & Garnier, H. E. (2009). Leadership for literacy coaching: The principal's role in launching a new coaching program. *Educational Administration Quarterly*, 45(5), 655–693.

- McDonald, M., Kazemi, E. & Kavanagh, S. S. (2013). Core Practices and Pedagogies of Teacher Education A Call for a Common Language and Collective Activity. *Journal of Teacher Education*, 64(5), 378–386.
- McGhee, M. W. & Lew, C. (2007). Leadership and writing: How principals' knowledge, beliefs, and interventions affect writing instruction in elementary and secondary schools. *Educational Administration Quarterly*, 43(3), 358–380.
- McGinnis, J. R., Parker, C., & Graeber, A. O. (2004). A cultural perspective of the induction of five reform-minded beginning mathematics and science teachers. *Journal of Research in Science Teaching*, 41(7), 720-747. doi: 10.1002/tea.20022
- Munter, C. (2014). Developing visions of high-quality mathematics instruction. *Journal for Research in Mathematics Education*, 45(5), 584–635.
- Munter, C. & Correnti, R. (2014). The Relation between visions of high-quality mathematics instruction and instructional improvement. In *Paper presented at the annual meeting of the National Council of Teachers of Mathematics*. New Orleans, LA.
- National Governors Association for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*.
- Naik, S. S. & Ball, D. L. (2014). Professional Development in a Laboratory Setting Examining Evolution in Teachers' Questioning and Participation. *Journal of Mathematics Education*, 7(2), 40–54.
- Nelson, B. S. (1997). Lenses on Learning: How Administrators' Ideas about Mathematics, Learning and Teaching Influence Their Approaches To Action in an Era of Reform. Center for the Development of Teaching Paper Series.

- Nelson, T. H., Slavit, D. & Deuel, A. (2012). Two dimensions of an inquiry stance toward student-learning data. *Teachers College Record*, 114(8), 1–42.
- Nelson, T. H., Lesseig, K., Slavit, D., Kennedy, C. & Seidel, R. (2015). Supporting Middle School Teachers' Implementation of STEM Design Challenges.
- Neumerski, C. M. (2013). Rethinking Instructional Leadership, a Review What Do We Know About Principal, Teacher, and Coach Instructional Leadership, and Where Should We Go From Here? *Educational Administration Quarterly*, 49(2), 310–347.
- Olson, J. & Barrett, J. (2004). Coaching teachers to implement mathematics reform recommendations. *Mathematics Teacher Education and Development*, 6, 73–91.
- Parrish, S. D. (2011). Number talks build numerical reasoning. *Teaching Children Mathematics*, 18(3), 198–206.
- Printy, S. M. (2008). Leadership for teacher learning: A community of practice perspective. *Educational Administration Quarterly*, 44(2), 187–226.
- Sassi, A., Bopardikar, A., Kimball, A. & Michaels, S. (2013). From “Sharing Out” to “Working Through Ideas:” Helping Teachers Transition to More Productive Science Talk. In *NARST*. Puerto Rico.
- Sassi, A. M., Morse, A. & Goldsmith, L. T. (1997). What Makes for a Good Beginning? Improvising in an Elementary Mathematics Teacher Inquiry Group.
- Schnellert, L. M., Butler, D. L. & Higginson, S. K. (2008). Co-constructors of data, co-constructors of meaning: Teacher professional development in an age of accountability. *Teaching and Teacher Education*, 24(3), 725–750.
- Senger, E. S. (1998). Reflective reform in mathematics: The recursive nature of teacher change. *Educational Studies in Mathematics*, 37(3), 199–221.

- Seymour, J. R. & Lehrer, R. (2006). Tracing the evolution of pedagogical content knowledge as the development of interanimated discourses. *The Journal of the Learning Sciences*, 15(4), 549–582.
- Simon, M. A. (1994). Learning mathematics and learning to teach: learning cycles in mathematics teacher education. *Educational Studies in Mathematics*, 26(1), 71–94.
- Slavit, D., Kennedy, A., Lean, Z., Nelson, T. H. & Deuel, A. (2011). Support for professional collaboration in middle school mathematics: A complex web. *Teacher Education Quarterly*, 113–131.
- Spillane, J. P., Diamond, J. B., Walker, L. J., Halverson, R. & Jita, L. (2001). Urban school leadership for elementary science instruction: Identifying and activating resources in an undervalued school subject. *Journal of Research in Science Teaching*, 38(8), 918–940.
- Stevens, R. & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. *Talking Mathematics in School: Studies of Teaching and Learning*, 107–149.
- Teacher Education by Design (TEDD). (n.d.). "Mathematics." Retrieved July 1, 2016 from the TEDD.org site: [http://tedd.org/?tedd\\_subject=mathematics](http://tedd.org/?tedd_subject=mathematics).
- Thompson, J., Windschitl, M. & Braaten, M. (2013). Developing a Theory of Ambitious Early-Career Teacher Practice. *American Educational Research Journal*, 50(3), 574–615.
- Van Es, E. A. (2011). Examining the development of a teacher learning community: The case of a video club. *Teaching and Teacher Education*.
- Visnovska, J. & Cobb, P. (2015). Learning about whole-class scaffolding from a teacher professional development study. *ZDM*, 1–13.

- Wilhelm, A. G. (2014). Mathematics teachers' enactment of cognitively demanding tasks: Investigating links to teachers' knowledge and conceptions. *Journal for Research in Mathematics Education*, 45(5), 636–674.
- Windschitl, M., Thompson, J., Braaten, M. & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*, 96(5), 878–903.
- Windschitl, M., Thompson, J., Braaten, M., Stroupe, D., Chew, C. & Wright, B. (2010). The beginner's repertoire: A core set of instructional practices for teacher preparation. In *annual meeting of the American Educational Research Association, Denver, CO*.
- Wood, D. (2007). Teachers' learning communities: catalyst for change or a new infrastructure for the status quo? *The Teachers College Record*, 109(3), 699–739.
- Yin, R. K. (2013). *Case study research: Design and methods – Third Edition*. Sage publications.
- Zech, L. K., Gause-Vega, C. L., Bray, M. H., Secules, T. & Goldman, S. R. (2000). Content-based collaborative inquiry: A professional development model for sustaining educational reform. *Educational Psychologist*, 35(3), 207–217.

## Appendix A

### Codebook of *A-Priori* Codes

Category	Sub category	Code	Description/Definition
School leadership			
	Instructional expectations	Aligned with ambitious vision	Principal expects teachers to use CMP2, to have students work in groups, looks for particular aspects of L-E-S lessons (launch, etc.)
		Vague or non-conflicting	Principal looks for things that aren't particularly ambitious but don't conflict with an ambitious vision either: objective posted, ss engaged
		Conflict with ambitious vision	Principal expects dedicated drill/procedures practice (e.g., test prep time included in every lesson)
		Data-focused	Teacher says principal is highly focused on teachers using student achievement data, potentially without specified purposes
	Supports for students	productive	School-wide supports for struggling students are aimed at supporting them to succeed in rigorous math instruction (VSMC rubrics)
		unproductive	School-wide supports for struggling students are disjointed from primary math instruction, focused primarily on test prep, potentially organized inequitably (e.g., by race or bubble kids)
		unknown quality	Teacher mentions supports (e.g., computer lab time, double dose class) but not enough to assess quality
	Instructional coaching	Coach expertise	Very Expert
On par with teachers			Coach VHQMI is no higher on average than the teachers s/he works with
Limited expertise			Coach has very little math expertise – e.g., 1s and 2s on VHQMI, 1SD below average on LMT test, perhaps has no math background
Coaching activities		high leverage	Teacher describes some combination of observations and feedback, coach modeling, co-teaching, co-planning, coaching cycle, working on math together

		low level	Teacher describes activities that are not about core instructional practice or connecting students/teaching/content: e.g., working with the coach to group students for Saturday school, or getting feedback on management
		Vague or tertiary	Teacher's description makes it difficult to assess quality: "we talk in the hall all the time, she's always there when I need her."
		infrequent	Teacher explicitly voices infrequency: "She maybe has been here one or two times all year, I'm not sure"
	Coach positioning	evaluative	Teacher says coach is there to evaluate instruction
		supportive	Teacher says coach is there to help/support/give advice
		vague	Teacher doesn't say, or doesn't know
		alternate coach	Flag marking the presence of more than 1 coach in the school
	Coach expectations	ambitious	In the case that the teacher describes / was asked to describe the coach's expectations for her work, these codes were used (and mirror those for school leaders)
		conflicting	
		vague	
		data	
PD	Math specific	Ambitious vision	PD described was organized around inquiry oriented curriculum or associated practices (leading discussions, launching tasks, etc.)
		Low level	PD was focused on strategies or tasks that sidestep instructional quality – e.g., electronic whiteboard techniques, or reverse engineering state test
		hard to tell	Teacher described math PD but didn't say enough about it to deduce the focus or quality
	Not math specific		PD described was not about math instruction in particular, but general instructional practices (e.g., grouping strategies, working w/ diverse students)
	T affect	positive	Teacher reported enjoying PD or finding it useful
		negative	Teacher did not enjoy the PD or was more critical than positive
TCT	Activities	High leverage	Teacher described engaging in TCT activities that involve connecting pts on the instructional triangle, such as modeling lessons or strategies, looking at student work together after teaching similar lessons, using the lesson planning protocol
		Low level (planning, pacing)	Teacher describes using TCT for efficient planning of weeks at a time; description indicates there's little focus on content or student thinking, but on coverage ('we talk about which sections we'll cover that week')

	Data analysis	Teacher describes using TCT to analyze student achievement data, with or without particular instructional purposes
	hard to tell	Teacher's description is too vague to differentiate between activity types
	creating and sharing materials	Usually low level, but not necessarily: using TCT to make assessments together, or find resources that can/should be used for particular standards
Facilitator	Expert Facilitator	The TCT facilitator has significant expertise in math instruction (IQA, VHQMI are 3s/4s; potentially MKT is high; VSMC productive)
	On par	TCT facilitator has expertise matching most teachers in the room
	Non math facilitator	Facilitator is not a math teacher or a coach with math expertise: e.g., data coach (who works with the whole building), principal, etc.
	No facilitator	TCT meetings don't have a set leader, teachers may take turns
Frequency	Monthly	TCT meetings occur monthly
	Every 2 weeks	“ “ every two weeks
	weekly	“ “ weekly
	more than weekly	“ “ twice or more per week
	3x month	“ “ three times per month
	informal	Teacher describes collaborating informally (“I go to her classroom between periods to talk about the lesson...”) either in place of or addition to formal meetings
affect	positive	Teacher describes TCT as helpful/useful/good use of time
	negative	Teacher is critical of TCT, doesn't find it a good use of time
Curriculum	High level	Primary curriculum used for math instruction (by the teacher) is inquiry-oriented, rigorous, usually CMP2 or materials from a Formative Assessment Lesson (FAL)
	Low level or skills based	Primary curriculum used for math instruction (by the teacher) is procedures-based, traditional demonstration-and-practice

## Appendix B

### Memo Template: Case Summary Form

Dissertation Pseudonym:

Case Type:

MIST years:

School leader(s):

Coach(es):

Notes about Expertise (trends or changes over time in VHQMI, VSMC, MKT, IQA):

#### School Leadership

*To what extent were school leaders at this teacher's school creating or shaping learning opportunities for this and other teacher? What were the teacher's perceptions of their leaders' expectations for their instruction?*

#### Coaching

*To what extent were math coaches at this teacher's school creating or shaping learning opportunities for this and other teacher? Did the teacher ever work one-on-one with the coach? What did they do together?*

#### PD

*To what extent did the teacher attend PD focused on ambitious math instruction – e.g., specific to implementing CMP tasks or using particular high-leverage practices? What was the teacher's perception of the helpfulness of these sessions?*

#### TCT

*What did the teacher report doing in typical TCT meetings? Who facilitated them? To what extent might the teacher's participation in these meetings shaped their practice? How?*

## Appendix C

### **Buy-in/Identification Memo Template**

Analytical memo re: (pseudonym) Case type: (I/D/S)

1. Over their years of improvement/decline, does the teacher seem to identify with or buy into the use of CMP as a primary material for math instruction in general and in their school?

2. [How] Does the teacher characterize the buy-in or sharedness of vision amongst other math teachers in his school? E.g., does the teacher sense any conflicting normative teaching identity among other teachers in the school?

3. Assess the strength of these claims based on the quality/amount of evidence in the teacher's interview transcripts. How much inference was needed to make these claims?

4. Open questions about the teacher:

## Appendix D

### Memo Template: Improver vs. Stable Teacher

Memo date:

*This memo should be used to document potentially critical differences between Improver (name) and Stayer (name) at school (school number).*

For each of the following questions, refer to the following a-priori categories of school supports for teacher learning (central to lit review):

- a. Curriculum
- b. TCT
- c. Coaching
- d. Instructional expectations/ P leadership
- e. PD

As well as the following aspects of individual teacher characteristics:

- g. Expertise: VHQMI, VSMC, MKT
  - h. Buy-in and identification with CMP and the district's reform efforts
- 

**1. What are the similarities and differences between the individual instructional expertise of the stayer and the improver?**

**2. Are there any differences or similarities in their buy-in or identification with the district's reform efforts? Cite evidence used to draw comparisons/contrasts.**

**3. Are there any differences between the stayer's and improver's engagement with school- or district-based supports for T learning? Provided there is adequate evidence to draw on, explain these differences in engagement, and cite the evidence used to do so.**

**4. Aside from the aspects of the school structure and individual expertise outlined above, were there any emergent themes of note which might situate the differneces between these two teachers' trajectories? cite evidence used.**

## Appendix E

### Triangulation Template 1: Year-specific buy-in/identification memo

Case (c\_code):

Case type (improver, decliner, or stable?):

Study year:

---

1. Based on your read of the interview transcript, to what extent does it seem the participant *identifies with* an ambitious vision of instruction? That is, to what extent does the participant view an ambitious vision of math instruction as their own personal goals for practice in their own classroom?

2. What evidence are you drawing on to make these conclusions? Are there any particular passages from the transcript that add strength to your argument?

3. What is the strength of this evidence? E.g., did the teacher narrate their own identification for you, or did you have to make a lot of inference?

4. Is there any portion of the interview transcript that troubles, or challenges your argument? Muddy water?

## Appendix F

### Template 2: Longitudinal buy-in/identification memo

Case (c\_code):

Case type (improver, decliner, or stable?):

Study years:

---

1. Based on your read of the interview transcript, to what extent does it seem the participant *identifies with* an ambitious vision of instruction? That is, to what extent does the participant view an ambitious vision of math instruction as their own personal goals for practice in their own classroom?

2. What is the strength of the evidence you're drawing on? E.g., did the teacher narrate their own identification for you, or did you have to make a lot of inference?

3. Were there any changes over time or specific shifts in this teacher's personal teaching identity over the years for which you read their interviews? Explain.

4. What is the strength of the evidence you're drawing on? E.g., did the teacher actually describe their own identification changing, or did you have to make inferences?

5. Is there any portion of the data that troubles, or challenges your argument? Muddy water?

## Appendix G

### Instructional expertise of case teachers

	Name		VHQMI	VSMC	MKT
improvers	Diane	6	Sophisticated/inquiry oriented	Mixed, improved over time	Pretty average all years
	Cynthia	8	Started low, improved over time	Mixed, prod sup	Algebra increased dramatically
	Anna	8	Never took a stance on task, RT and discourse scores came up over time	unproductive/mixed, improved once full-time coach	Average
	Mario	6	Very ambitious in y4 (his first year of teaching), briefly gave up, then came back to more 'transitional' after that – described procedures with connections as his goal (“processing”)	Productive/mixed	Below average
decliners	Karyn	6	Started fairly ambitious (save for MT), declined	Unproductive	fairly high
	Alishia	8	Started fairly traditional, improved over time to form view (3/3/3)	Mixed/unproductive	Slightly above average, improved over time
	Li	6	Pretty procedural, lots of emphasis on monitoring/behavior management	Very unproductive	Above average, esp algebra
	Penny	6	Fairly high/inquiry oriented	Very unproductive	Average, low in algebra
Stable teachers	Lois	7	Traditional but wants discourse, no stance on tasks	Mixed/unproductive	High
	Roger	6	Pronouncedly traditional	Unproductive	Average
	Roxanne	8	No stance on tasks, sees teacher as monitor, but wants discourse	Unproductive	Low/avg (oddly not higher in algebra, which she taught)

## Appendix H

### Expertise of case teachers' school leader(s)

District	School (Case teachers)	Instructional leaders (P = principal, AP = assistant principal)	Years they were leaders	Notes on VHQMI
D	Cypress (Diane, Cathy, and Roger)	R (P) and A (AP) L (P)	1-4; 5-7	Both R and A got to be pretty sophisticated in years 3-4. L was fairly traditional VHQMI at first, but had developed a fairly sophisticated vision by y7.
D	Poplar (Karyn)	S (P) 5-6 J (P) 7	5-7	S and J both had a pretty traditional VHQMI. J's wasn't able to be coded.
B	Two Lagoons (Mario, Lois)	D(P) & Vera (AP) (y4) J(P)	1-4 5-7	D and Vera both had fairly ambitious vision of instruction – Vera esp. J's VHQMI was pretty traditional in year 5 but improved to more transitional/ambitious in y6-7.
B	Park Falls (Anna, Roxanne)	L (P)	2-4	In general, fairly traditional VHQMI, though in year 3 her vision seemed more sophisticated (role of the teacher and discourse), but then went back down in y4
B	Salt Marsh (Li)	J (P)	5-7	J's vision was pretty traditional in y6-7, though he did want to see a mix of small-group and whole group discourse, he never really had anything specific he was looking for in math tasks, and nothing in role of the teacher beyond monitoring students being on task
B	Meadowbrook (Penny)	C (P)	1-3	Started out quite traditional, but really developed quite sophisticated VHQMI by year 3 (3s and 4s! first in discourse, then in RT)