Educational Design Research to Support System-Wide Instructional Improvement

Erin Henrick and Paul Cobb

Vanderbilt University

Department of Teaching and Learning

1930 South Drive, 240 Wyatt Center

Peabody Box 330

Nashville, TN 37203

615.322.7568 (office)

615.322.8999 (fax)

erin.henrick@vanderbilt.edu

paul.cobb@vanderbilt.edu

Kara Jackson

McGill University

Faculty of Education

Department of Integrated Studies in Education

Montreal, QC

Canada H3A 1Y2

kara.jackson@mcgill.ca

Abstract

In this chapter, we describe a methodology for conducting educational design research to support system-wide instructional improvement in mathematics and draw on one of the few design studies that does this as an illustrative case. Design studies conducted at the level of an educational system are interventionist in nature, and can address both the complexity of educational settings and the problems that educational system leaders, school leaders, and teachers encounter as they work to improve the quality of classroom instruction, school instructional leadership, and ultimately, students' mathematics learning. This chapter describes the theoretical background for this approach, in which the issue of what it takes to support instructional improvement on a large scale is framed as an explicit focus of empirical investigation.

Our purpose in this chapter is to describe a methodology for conducting educational design research to support large scale instructional improvement in mathematics. In countries with centralized education systems, large scale might mean instructional improvement at the national level. In countries with decentralized education systems, the appropriate organizational unit is the largest administrative jurisdiction that can support coordinated improvement efforts. In this chapter, this largest unit will be referred to as the *educational system* or *system*.

For the purpose of this chapter, we define design research as a family of methodological approaches in which research and the design of supports for learning are interdependent. On the one hand, the design of supports serves as the context for research and, on the other hand, ongoing and retrospective analyses are conducted in order to inform the improvement of the design (Gravemeijer, 1994; Schoenfeld, 2006). Design research methodology has become increasingly prominent in the learning sciences and in several related fields of educational research including mathematics education in recent years. Most design research studies focus on students' mathematical learning either as they interact one-on-one with a researcher (e.g., Cobb & Steffe, 1983; Lobato, 2003; Steffe & Thompson, 2000) or as they participate in classroom processes (e.g., Cobb, Confrey, diSessa, et al., 2003; Design-Based Research Collaborative, 2003). In comparison, design studies conducted to support and investigate teachers' learning are far less common, and design studies conducted to study the process of supporting improvements in the quality of mathematics teaching on a large scale have, until recently, been extremely rare. As a consequence, there are currently few examples of design research studies that have been conducted at the level of an educational system, or that focus simultaneously on teachers' development of instructional practices and the school and system settings in which they develop and refine those practices (e.g., Cobb, Zhao, & Dean, 2009; Cobb, McClain, Lamberg, & Dean, 2003; Fishman, Marx, Blumenfeld, & Krajcik, 2004). However, design research at scale (Cobb & Jackson, 2012) and closely related approaches such as design based implementation research (Penuel, Fishman, Cheng, & Sabelli, 2011) and improvement science research (Bryk, 2009) are gaining momentum.

As Stein (2004) observed, research in mathematics education has not, to this point, investigated how the school and system settings in which mathematics teachers work can be organized to support their ongoing learning. Key aspects of these settings include the materials and associated resources that teachers use as a basis for their instruction, the formal and informal sources of assistance on which they can draw, as well as to whom and for what they are accountable. Design studies conducted at the level of an educational system is interventionist in nature, and can address both the complexity of these educational settings and the problems that educational system leaders, school leaders, and teachers encounter as they work to support improvements in the quality of classroom instruction, school instructional leadership, and ultimately, students' mathematics learning.

We illustrate the methodology for investigating and supporting system-wide instructional improvement by framing one of the few design studies of this

type as a sample case. The study, *Designing Learning Organizations for Instructional Improvement in Mathematics* (known as MIST)*,* was conducted in the United States and investigated how school- and system-level supports and accountability relations impacted the quality of mathematics instruction in middle-grades schools that served students aged 12-14. The study was a four-year collaboration with district leaders, school leaders, and mathematics teachers in four city-wide school systems that served a total of 360,000 students. Before describing and illustrating the methodology, we first provide background information on the United States educational context. We also detail the vision of high-quality mathematics instruction that oriented our research agenda and recruitment of participating school systems.

The United States Context

The United States educational system is decentralized, and there is a long history of the local control of schooling. Each U.S. state is divided into a number of independent school districts. In rural areas, many districts serve less than 1,000 students whereas a number of urban districts serve more than 100,000 students. In the context of the U.S. educational system, urban districts are the largest jurisdictions in which it is feasible to design for improvement in the quality of instruction (Supovitz, 2006).

The federal government's role in the educational system in the U.S. has increased significantly in recent years following the passing of the No Child Left Behind Act (NCLB) in 2001. States receive incentives to set standards for students' mathematics achievement, develop standardized assessments aligned with the standards, and implement accountability measures to promote increases in achievement for all students and for specific sub-groups (e.g., racial and ethnic categories, socio-economic status, students who receive special education services). Districts and schools are sanctioned if they fail to meet goals for "adequate yearly progress" (AYP) on state assessments.

As a result, school districts are under great pressure to improve student achievement in mathematics. In addition to responding to accountability pressures, urban school districts in the United States face a number of other challenges that impact improvement initiatives. These challenges include limited financial resources, under-prepared teachers, and high teacher turnover (Darling-Hammond, 2007).

Unfortunately, most U.S. school districts do not have the capacity to respond to these accountability demands in a productive manner (Elmore, 2006). Many districts are implementing short-term interventions aimed at "teaching to the test," and some are attempting to game the assessment system (Heilig & Darling-Hammond, 2008). In addition, districts frequently expend considerable resources on different (and even conflicting) improvement policies, abandoning each for the next when student achievement does not improve quickly, without understanding

the challenges of implementing particular policies. This policy churn (Hess, 1999) can cause frustration for teachers and does not help the larger educational community understand how improvement in student achievement can be supported at the scale of a large school district.

A minority of districts is responding to accountability demands by attempting to improve the quality of classroom instruction. These districts are attempting to support teachers' development of high quality instructional practices that will ultimately lead to improvement in student achievement (Elmore, 2004). Concurrently, the role of the principal is shifting from school manager to instructional leader, with an increased responsibility to support instructional reforms in each content area (Nelson & Sassi, 2005; Fink & Resnick, 2001). To date, efforts to support fundamental improvements in teachers' instructional practices on a large-scale have rarely been successful, and there are no proven models regarding how this can be accomplished (Elmore, 2004; Gamoran et al., 2003). Furthermore, although research on mathematics teaching and learning has made significant advances in recent years, these advances have had limited impact on the quality of instruction in most U.S. classrooms. In addition, research in both mathematics education and in educational policy and leadership can provide only limited guidance to districts attempting to respond to high stakes accountability pressures by improving the quality of mathematics instruction.

An Orienting Vision of High-Quality Mathematics Instruction

The four urban school systems, or districts, that we recruited for the MIST study were all pursuing similar agendas for instructional improvement in mathematics. These agendas were oriented by goals for students' mathematics learning that are relatively ambitious in the U.S. context. These system-level goals emphasized students' development of conceptual understanding as well as procedural fluency in a range of mathematical domains, students' use of multiple representations, students' engagement in mathematical argumentation to communicate mathematical ideas effectively, and students' development of productive dispositions towards mathematics (U.S. Department of Education, 2008; Kilpatrick, Swafford & Findell, 2001; National Council of Teachers of Mathematics, 2000). These student learning goals in turn oriented leaders of the four collaborating districts as they specified high-quality mathematics instructional practices that could be justified in terms of student learning opportunities (Kazemi, Franke, Lampert, 2009). The resulting view of high-quality instruction has been referred to in the U.S. as *ambitious teaching* (Lampert & Graziani, 2009; Lampert, Beasley, Ghousseini, Kazemi & Franke, 2010).

Ambitious teaching requires teachers to build on students' solutions to challenging tasks while holding students accountable to learning goals (Kazemi et al., 2009). Recent research in mathematics education has begun to delineate a set

of high-leverage instructional practices that support students' achievement of ambitious learning goals (Franke, Kazemi & Battey, 2007; NCTM, 2000). These practices include launching challenging tasks so that all students can engage substantially without reducing the cognitive demand of tasks (Jackson, Garrison, Wilson, Gibbons, & Shahan, in press), monitoring the range of solutions that students are producing as they work on tasks individually or in small groups (Horn, 2012), and building on these solutions during a concluding whole-class discussion by pressing students to justify their reasoning and to make connections between their own and others' solutions (Staples, 2007; Stein, Engle, Smith, & Hughes, 2008). These practices differ significantly from the current practices of most U.S. teachers, and their development involves reorganizing rather then merely adjusting and elaborating current practices. The learning demands for teachers include developing a deep understanding both of the mathematics on which instruction focuses and of students' learning in particular mathematical domains. In addition, it involves developing the new high-leverage instructional practices outlined above (e.g., launching cognitively-demanding tasks effectively; orchestrating whole class discussions of students' solutions that focus on central mathematical ideas).

The agenda for instructional improvement that the four collaborating school systems were pursuing is specific the U.S. context and was influenced by the recommendations of several professional organizations including the National Council of Teachers of Mathematics (1989, 2000) and the Common Core State Standards Initiative (2010). Improvement efforts in other countries might be oriented by a different vision of high-quality mathematics instruction.   The methodology that we describe will nonetheless be relevant to all cases where instructional improvement goals involves significant teacher learning and requires teachers to reorganize rather than merely elaborate their current classroom practices.

In the remainder of this chapter, we describe the key aspects of design studies conducted to investigate and support system-wide improvement in mathematics instruction. Although we draw on the MIST study to clarify the rationale for certain tools and processes, our intent is to describe the methodology in broad terms.

Design Studies to Investigate and Support System-Wide Improvement in Mathematics Instruction

The overall goal of design research at the level of an education system is to investigate what it takes to support instructional improvement at scale (Byrk & Gomez, 2008, Coburn & Stein 2010, Roderick et al., 2009) by testing and revising conjectures about school- and system-level supports and accountability relations. Design studies of this type aim to both support and investigate the *process* of instructional improvement at scale by documenting 1) the trajectories of (interrelated) changes in the school- and system-level settings in which mathematics teach-

ers work, their instructional practices, and their students' learning, and 2) the specific means by which these changes are supported and organized across the system (Cobb & Smith, 2008).

Design studies of this type have two primary objectives. The first objective is pragmatic, and is to provide leaders of the collaborating educational systems with timely feedback about how their improvement strategies or policies are actually playing out that can inform the ongoing revision of instructional improvement efforts. The second objective is theoretical, and is to contribute to the development of a generalizable theory of action (Argyris and Schön, 1974) for system-wide instructional improvement in mathematics by synthesizing findings across multiple educational systems.

Design studies conducted at any level involve iterative cycles of designing to support learning and of conducting analyses that inform the revision of the current design. In contrast to studies conducted to investigate students' learning, design studies at the system level necessarily entail a partnership with system leaders. As a consequence, cycles at this level also include a feedback phase in which researchers share findings with system leaders who have the ultimate authority for making decisions about improvement strategies. The length of the cycles is much longer that in other types of design research. For example, in the MIST study, each cycle spanned an entire school year.

In the sections below, we describe the following aspects of the methodology:

1. developing an initial set of conjectures that comprise an initial theory of action about school- and system-level supports and accountability relations;
2. recruiting collaborating educational systems;
3. employing an interpretative framework for assessing an educational system's designed and implemented instructional improvement strategies;
4. conducting successive design, analysis and feedback cycles by: a) documenting each collaborating system's current improvement strategies, b) collecting and analyzing data on how those strategies are actually playing out, c) sharing findings and recommendations system leaders in time to inform their revision of improvement plans, and d) assessing the influence of recommendations on the collaborating system's instructional improvement strategies;
5. testing and revising conjectures that comprise a theory of action for system-wide instructional improvement based on ongoing feedback analyses, the current research literature, and retrospective analyses of data collected in successive cycles.

Developing Initial Conjectures

The basic goal of a design study conducted at any level is to improve an initial design for supporting learning by testing and revising conjectures inherent in the design about the course of participants' learning and the means of supporting their learning (Cobb et al., 2003). A key concern when preparing for a system-level design study is therefore to develop an initial set of conjectures for what it would take to support improvement in the quality of mathematics teaching across an entire system.

In the MIST study, we found it valuable to follow the basic tenets of design as articulated by Wiggins and McTighe (1998) and develop initial conjectures by mapping out from the classroom (cf. Elmore, 1979-80). The first step in the process is to specify explicit goals for students' mathematical learning and an associated research-based vision of high-quality mathematics instruction. The learning demands for teachers can then be identified by comparing the vision of high-quality mathematics instruction that constitutes the goal for teachers' learning with their current instructional practices.

The second step is to develop an initial, tentative, and eminently revisable theory of action by formulating conjectures about both supports for teachers' learning and accountability relations that press them to improve their practices. These conjectures should clearly attend to teacher professional development and to instructional materials and associated tools designed for teachers to use. However, it also proved important in the MIST study to broaden our purview by considering other types of possible support such as mathematics teacher collaborative meetings scheduled during the school day, the colleagues to whom teachers turned for instructional advice during the school day, and mathematics teacher leaders or coaches who were charged with supporting teachers in their classrooms and during collaborative meetings. In addition, research on school instructional leadership oriented us to consider the role of principals and other school leaders in pressing and holding teachers accountability for improving the quality of instruction.

It is important to note that conjectures about supports and accountability relations for teachers' learning typically have implications for the practices of members of other role groups. For example, conjectures about the role of coaches in supporting teachers' learning have implications for the practices of system leaders responsible for hiring coaches and for supporting their development of effective coaching practices. Similarly, conjectures about school leaders' role in communicating appropriate instructional expectations to teachers have implications for the practices of others in the system who are charged with supporting them in deepening their understanding of high-quality mathematics instruction.

In following this process of mapping out from the classroom in the MIST study, it proved critical to balance the ideal with the feasible by taking account of each collaborating system's current capacity to support members of different role groups in improving their practices. As we worked through this process of formu-

lating initial conjectures, we also found that the challenge of improving classroom instruction had implications for the practices of personnel at the highest levels of the four collaborating systems. As a consequence, it proved essential to formulate testable conjectures about the means of supporting the learning of mathematics teachers, mathematics coaches, school leaders, and system leaders in a coordinated manner. It also became apparent as we worked through this process that issues of mathematical content really matter. The mathematical learning goals for students have direct implications for the vision of high-quality instruction and thus for the learning demands on the teachers. These learning demands in turn have implications for conjectures about supports and accountability relations for teachers' learning, and thus for the practices of personnel at all levels of the system.

Research on instructional improvement at the level of an educational system is thin, and gets thinner the further one moves away from the classroom. In order to formulate MIST conjectures about potentially productive school- and system-level supports, we drew on the limited number of relevant empirical studies and conceptual analyses available in the mathematics education literature on mathematics teaching, professional development, and teacher collaboration (Kilpatrick, Martin, & Schifter, 2003, Cobb & McClain, 2001; Franke & Kazemi, 2001; Gamoran, Secada, & Marrett, 2000; Kazemi & Franke, 2004; Little, 2002; Stein, Silver, & Smith, 1998; Coburn and Russell, 2008) and the literature on education policy and leadership that viewed policy implementation as involving learning (Blumenfeld, Fishman, Krajcik, Marx, & Soloway, 2000; Coburn, 2003; McLaughlin & Mitra, 2004; Stein, 2004; Tyack & Tobin, 1995).

The resulting conjectures specified school and district structures, social relationships, and material resources that we anticipated might support mathematics teachers' and instructional leaders' ongoing learning. These conjectures assumed that the district has adopted research-based, inquiry-oriented mathematics textbooks and would provide sustained teacher professional development.

One of our conjectures drew on research in educational policy and leadership with indicated the importance of teachers and school leaders having a common improvement agenda. We therefore conjectured that a shared vision of high-quality mathematics instruction in schools would be associated with instructional improvement. A second conjecture specified that instructional improvement will be supported both if school instructional leaders support and hold mathematics teachers accountable for developing high-quality instructional practices, and if district leaders hold school leaders accountable for assisting mathematics teachers in improving their instructional practices. We also conjectured that greater improvements in overall student mathematics achievement would occur if school mathematics programs were de-tracked so that classes were heterogeneous rather than organized by current student achievement. This conjecture drew on research that indicated that "tracking does not substantially benefit high achievers and tends to put low achievers at a serious disadvantage" (Darling-Hammond, 2007, p. 324). A fourth conjecture indicated the importance that we attributed to the alignment of goals and strategies for instructional improvement across district

central office units, particularly Leadership, the department responsible for supporting and holding school leaders accountable, and Curriculum and Instruction, the department responsible for selecting instructional materials and for providing professional development for teacher and coaches. Cobb and Smith (2008) provide more detail on these initial conjectures.

Recruiting Collaborating Educational Systems

When preparing for a system-level design study, it is essential to formulate explicit criteria for selecting educational systems to target for participation in the study. In doing so, it is important to remember that it is system leaders rather than researchers who have the ultimate authority to determine both goals for students' mathematics learning and what counts as high-quality mathematics instruction. One important selection criterion for any system-level design study is therefore that system leaders' views of high-quality mathematics instruction are similar to those of the researchers.

In the case of MIST, a second important criterion was that the district be typical of large urban districts in the US in terms of persistent patterns of low student achievement (including disparities in achievement between historically disadvantaged and advantaged groups of students); high teacher turnover; and relatively large numbers of novice teachers.

A third selection criterion related to how a district was responding to accountability demands. Given our focus on instructional improvement at scale, we sought to recruit districts that were among the minority that were responding by focusing on the quality of classroom mathematics instruction and students' mathematical learning. In this regard, all of the districts with which we collaborated were atypical of urban districts in that they sought to improve student achievement in the middle grades mathematics by supporting teachers' development of ambitious mathematics teaching practices of the type described earlier in the paper. The collaborating districts' goals for students' mathematical learning extended beyond improving achievement on state tests, and included a concern for students gaining admission to college and succeeding when they get there.

A fourth selection criterion was that the districts framed the problem of instructional improvement in terms of teacher learning and were attempting to implement a reasonably well worked out set of improvement strategies (Elmore, 2006). Each of the four district's pre-existing improvement strategies for high-quality mathematics instruction aligned with current research on mathematical learning, and encompassed curriculum, teacher professional development, and school instructional leadership. Examples of such strategies include adopting an inquiry-oriented textbook series for middle-grades mathematics, providing high quality teacher professional development, scheduling time during the school day for mathematics teachers to collaborate, recruiting and supporting a cadre of

school- or district-based mathematics coaches, and supporting instructional leaders' development of instructional leadership practices through professional development..

A fifth criterion was that middle-grades mathematics was a priority area for the districts. The four districts were committed to providing time and resources to further their instructional improvement efforts in middle-grades mathematics, and considered that participating in our study could contribute to these efforts.

Not surprisingly, the number of urban districts that met our five criteria was limited. We identified three of the districts with the assistance of the Institute for Learning, a national organization that partners with districts to guide their development and implementation of improvement policies.

In conducting a design study of instructional improvement at scale, it is typically not feasible to collect data in all schools in the systems that have been recruited. Given the intent of a study of this type, we recommend purposefully selecting schools that reflect the overall variation in student performance and capacity for instructional improvement across all schools in each system. Teachers might then be recruited randomly within schools, or they might be selected purposefully to reflect variation in quality of current instructional practices.

In MIST, we recruited 30 middle-grades mathematics teachers from between six and ten schools in each of the four districts, together with 20 school and district leaders in each district. We found that teachers often agreed to participate in the MIST study because they saw it as an opportunity to have their perspective taken into account when system leaders formulated district improvement policies for middle-grades mathematics.

Using an Interpretive Framework to Assess Designed and Implemented Improvement Strategies

Based on our experience in the MIST study, two types of conceptual tools are important when conducting investigations of this type. The first tool is a theory of action for large-scale instructional improvement in mathematics that consists of testable conjectures about supports and accountability relations. As we described above, we developed an initial theory of action in the MIST study by drawing on then current research in mathematics education and the learning sciences, educational leadership, and educational policy before we began working in the four collaborating systems. We tested and revised the conjectures that comprised this initial theory of action as we conducted successive cycles of design and analysis in the course of the study. The theory of action is central to the design phase of the iterative design and analysis cycles that characterize design research at the system level.

The second type of conceptual tool is an interpretive framework that can be used to assess the potential of the collaborating systems' designed or intended strategies to contribute to instructional improvement. This tool is central to the analysis phase of each cycle. During the first two years of the MIST study, we developed an interpretive framework that distinguishes between four general types of supports: new positions, learning events (including professional development), organizational routines, and tools. These types of supports capture all the improvement strategies that the four collaborating systems attempted to implement across the four years. In developing the framework, we drew on research in the learning sciences, teacher learning, and related fields to assess the potential of each general type of support to scaffold teachers', coaches', and school leaders' reorganization of their practices.

We clarify the nature of each type of support and its potential to support practitioners' learning in the following paragraphs. As will become apparent, the framework reflects the view that co-participation with others who have already developed relatively accomplished practices is essential when the learning demands of an improvement strategy require the reorganization rather than the extension or elaboration of current practices (Lave & Wenger, 1991; Rogoff, 1997; Sfard, 2008).

New Positions

School- and system-level strategies for instructional improvement typically include changes in the responsibilities of existing positions, such as principals becoming effective instructional leaders in mathematics. In addition, improvement efforts often include the creation of new positions whose responsibilities include supporting others' learning by providing expert guidance. For example, an educational system might create the position of a school-based mathematics coach in each school, whose responsibilities include supporting their principals in becoming instructional leaders in mathematics. This improvement strategy assumes that the coaches have developed greater expertise as instructional leaders in mathematics and can therefore guide principals as they attempt to support mathematics teachers' improvement of their classroom practices (Bryk, 2009; Spillane & Thompson, 1997).

The importance that we attributed to the expertise or knowledge-in-practice of the holder of the new position follows directly from Vygotskian accounts of human development (Kozulin, 1990; van der Veer & Valsiner, 1991; Vygotsky, 1978) and is supported by studies of apprenticeship and coaching (Brown, Collins, & Duguid, 1989). We therefore view the provision of expert guidance by creating new positions (or changing the responsibilities of existing positions) as a primary support for learning. The extent to which the investment in the new position will pay off is likely to be influenced by a variety of factors in addition to the expertise of the appointee. These additional factors include the

overall coherence of instructional improvement strategies and the extent to which the expert and target of policy co-participate in activities that are close to the intended forms of practice.

Learning Events

Large-scale instructional improvement efforts typically include professional development for teachers and, on occasion, for members of other role groups including principals. We view professional development sessions as instances of learning events, which we define as scheduled meetings that can give rise to opportunities for targets of policy to improve their practices in ways that further policy goals. We take account of both learning events that are intentionally designed to support targets' learning and those that might give rise to incidental learning.

Intentional learning events

A distinction that proved useful in the MIST study when analyzing the strengths and weaknesses of improvement strategies is that between intentional learning events that are ongoing and those that are discrete. The two key characteristics of *ongoing intentional learning events* are that they are designed as a series of meetings that build on one another, and that they involve a relatively small number of participants. As an example, a mathematics specialist might work regularly with middle-school principals as a group in order to support them in recognizing high-quality mathematics instruction when they make classroom observations. Because a small number of participants is involved, the group might evolve into a genuine community of practice that works together for the explicit purpose of improving their practices.

It is important to note that although communities of practice can be productive contexts for professional learning (Horn, 2005; Kazemi & Hubbard, 2008), the emergence of a community of practice does not guarantee the occurrence of learning opportunities that further policy goals (Bryk, 2009). Recent research in both teacher education and educational leadership indicates the importance of interactions among community members that focus consistently on issues central to practice (Marks & Louis, 1997) and that penetrate beneath surface aspects of practice to address core suppositions, assumptions, and principles (Coburn & Russell, 2008). This in turn suggests the value of one or more members of the community having already developed relatively accomplished practices so that they can both push interactions to greater depth (Coburn & Russell, 2008) and provide concrete illustrations that ground exchanges (Penuel, et al., 2006). The critical role of expertise in a community of practice whose mission is to support participants' learning is consistent with the importance attributed to "more

knowledgeable others" in sociocultural accounts of learning (Bruner, 1987; Cole, 1996; Forman, 2003).

The key aspects of ongoing intentional learning events that we have highlighted are consistent with the qualities of effective teacher professional development identified in both qualitative and quantitative studies. These qualities include extended duration, collective participation, active learning opportunities, a focus on problems and issues that are close to practice, and attention to the use of tools that are integral to practice (Borko, 2004; Cohen & Hill, 2000; Desimone, Porter, Garet, Suk Yoon, & Birman, 2002; Garet, Porter, Desimone, Birman, & Yoon, 2001). We view ongoing intentional learning events that have these qualities as a primary means of supporting consequential professional learning that involves the reorganization of practice.

Discrete intentional learning events include one-off professional development sessions as well as a series of meetings that are not designed to build on each other. For example, system leaders might organize monthly meetings for principals. These meetings would be discrete rather than ongoing intentional learning events if principals engage in activities that focused on instructional leadership in mathematics only occasionally, and these activities do not build on each other. Discrete intentional learning events can be valuable in supporting the development of specific capabilities that elaborate or extend current practices (e.g., introducing a classroom observation tool that fits with principals' current practices and is designed to make their observations more systematic). However, they are by themselves unlikely to be sufficient in supporting the significant reorganization of practice called for in systems that are pursuing ambitious instructional agendas.

Incidental learning events

Learning opportunities are not limited to those that are intentionally designed, but can also arise incidentally for targets of policy as they collaborate with others to carry out functions of the school or educational system. For example, if principals meet regularly with mathematics coaches to discuss the quality of mathematics teaching in the school, these meetings could provide learning opportunities for the principal even though these meetings were not designed to support the principals' learning. In general, the extent to which regularly scheduled meetings with a more knowledgeable other involve significant learning opportunities depends on both the focus of interactions (e.g., the nature of teachers' classroom practices and student learning opportunities) and on whether the expert has in fact developed relatively accomplished practices and the novice recognizes and defers to that expertise (Elmore, 2006; Mangin, 2007). However, the strategy of relying primarily on incidental learning events to support professional learning appears to be extremely risky.

New organizational routines

In addition to creating new positions and planning learning events, instructional improvement policies sometimes include the specification of new organizational routines. Feldman and Pentland (2003) define organizational routines as "repetitive, recognizable patterns of interdependent actions, carried out by multiple actors" (p. 94). Investigations of organizational routines in school settings demonstrate that they can play a critical role in ensuring continuity and thus school stability in the face of high staff turnover (Spillane, Mesler, Croegaert, & Sherer, 2007). In addition, these studies clarify that organizational routines often evolve incrementally in the course of repeated enactments and can therefore also be a source of organizational flexibility (Feldman, 2000, 2004). Furthermore, as Sherer and Spillane (2011) illustrate, the introduction of carefully designed organizational routines can be an important means of supporting learning.

An illustration of an organizational routine would be system leaders expecting principals to conduct Learning Walks™ with the mathematics coach at their schools on a regular basis. A Learning Walk™ is a repetitive, recognizable pattern of actions that involves determining the focus of classroom observations (e.g., the extent to which teachers maintain the cognitive challenge of tasks throughout the lesson), selecting classrooms to visit, observing a classroom, and then conferring to discuss observations before moving on to the next classroom. In addition, a Learning Walk™ is carried out by multiple actors, namely the principal, mathematics coach, and the observed teachers. This organizational routine provides opportunities for the mathematics coach to support the principal in coming to recognize key aspects of high-quality mathematics instruction.

In this example, the organizational routine is conducted independently of any formally scheduled meetings. Other organizational routines might be enacted during either intentional or incidental learning events. For example, a mathematics specialist working with a group of principals might introduce an organizational routine that first involves having principals collect student work on the same instructional task from one or more classrooms in their schools, next having the principals analyze the quality of the student work in small groups, and finally pressing the principals to delineate the characteristics of high-quality work during a subsequent whole group discussion. We consider organizational routines in which a more knowledgeable other scaffolds relative novices' learning as they co-participate in a sequence of activities that are close to practice to be a potentially productive means of supporting professional learning (Grossman & McDonald, 2008; Lampert & Graziani, 2010).

New Tools

In speaking of tools, we refer to material entities that are used instrumentally to achieve a goal or purpose. Work in the learning sciences and in teacher

professional development indicates that introducing carefully designed tools is a primary means of supporting learning (Borko, 2004; Cobb, Zhao, & Dean, 2009; Lehrer & Lesh, 2003; Meira, 1998). In the context of large-scale instructional improvement efforts, designed tools can also play a second important role by supporting members of a particular role group in developing compatible practices, and by supporting the alignment of the practices developed by members of different role groups (e.g., teachers, principals, coaches). Examples include textbooks, curriculum guides, state mathematics objectives, classroom observation protocols, reports of test scores, student written work, and written statements of school and educational system policies.

Large-scale instructional improvement efforts almost invariably involve the introduction of a range of new tools designed to be used in practice, including newly adopted instructional materials and revised curriculum frameworks for teachers, and new classroom observation protocols and data management systems for principals. The findings of a number of studies conducted in the learning sciences substantiate Pea's (1993) claim that the incorporation of a new tool into current practices can support the reorganization of those practices (Lehrer & Schauble, 2004; Meira, 1998; Stephan, Bowers, & Cobb, 2003). However, it is also apparent that people frequently use new tools in ways that fit with current practices rather than reorganizing those practices as the designers of the tool intended (Wenger, 1998). For example, the findings of a number of studies of policy implementation and of teaching indicate that teachers often assimilate new instructional materials to their current instructional practices rather than reorganize how they teach as envisioned by the developers of the materials (Cohen & Hill, 2000; Remillard, 2005; Spillane, 1999). These findings suggest that the design of tools for professional learning should be coordinated with the development of supports for their increasingly accomplished use.

As a first design heuristic, it is important that users see a need for the tool when it is introduced (Cobb, 2002; Lehrer, Schauble, & Penner, 2000). This implies that either the tool should be designed to address a problem of current practice or it should be feasible to cultivate the need for the tool during intentional learning events. As an illustration, consider a classroom observation protocol that has been designed to support principals in focusing not merely on whether students are engaged but also on whether significant learning opportunities arise for them. Most principals are unlikely to see a need for the new observation form unless it is introduced during a series of intentional learning events that might, for example, focus on the relation between classroom learning opportunities and student achievement.

Second, it is also important that the tool be designed so that intended users can begin to use it shortly after it has been introduced in relatively elementary ways that are nonetheless compatible with the designers' intentions and do not involve what A. Brown (1992) termed lethal mutations. In the case of our example, it would seem advisable to minimize the complexity of the observation protocol given the significant reorganization of practice that most principals would have to

make to use it in a way compatible with the designers' intentions (Nelson & Sassi, 2005).

Third, in using the tool in rudimentary but intended ways, users begin to reorganize their practices as they incorporate the tool. The challenge is then to support their continued reorganization of practice by scaffolding their increasingly proficient use of the tool either during intentional learning events or as they co-participate in organizational routines with an accomplished user (J. S. Brown & Duguid, 1991; Lave, 1993; Rogoff, 1990). In the case of the observation protocol, for example, mathematics coaches might support principals' use of the tool as they conduct Learning Walks™ together. Just as the failure to provide sustained teacher professional development around a new curriculum can lead to difficulties (Crockett, 2007), failure to scaffold principals', coaches', and others' use of new tools is also likely be problematic.

Summary

Our analysis of the four types of support for learning indicates that improvement strategies that are likely to be effective in supporting consequential professional learning involve some combination of new positions that provide expert guidance, ongoing intentional learning events in which tools are used to bridge to practice, carefully designed organizational routines carried out with a more knowledgeable other, and the use of new tools whose incorporation into practice is supported. We do not discount the support that discrete intentional learning events and incidental learning events might provide and recommend taking them into account when assessing systems' improvement strategies. However, research on professional learning and on students' learning in particular content domains indicated that they are, by themselves, rarely be sufficient to support significant reorganizations of practice (Garet, et al., 2001). The analysis we conducted during the MIST study of the four districts' instructional improvement efforts over a four-year period is consistent with this conclusion.

Conducting Design, Analysis and Feedback Cycles

Thus far, we have discussed the key issues that need to be addressed when preparing for a system-level design study. We now focus on the process of conducting a study by enacting successive design, analysis, and feedback cycles. Each of the four cycles we conducted in the MIST study spanned an entire school year, which is much longer than in other types of design experiments (a day in the case of a classroom design study and a few weeks or less for a professional development study).

In planning cycles, it is important to take account of patterns in system leaders' work across the school year. In the U.S. educational systems, the school year runs from August until May or the beginning of June. In the MIST study, we delayed interviewing district leaders to learn about their current instructional improvement plans until October of each year after they had finalized their plans for that school year. We then determined that January-March would be the best time to collect data because it would give us enough time to conduct the feedback analyses, and would not interfere with standardized testing, which typically occurs near the end of the school year. We shared our feedback and recommendations with district leaders in May of each year so they could take account of our findings when they revised district instructional improvement strategies over the summer.

A) Documenting Current Instructional Improvement Strategies

The first phase of a cycle involves documenting the vision of high-quality mathematics instruction that orients each collaborating system's instructional improvement initiative and the strategies that each system is implementing in an attempt to achieve its vision. In the MIST study, it proved feasible to document the four collaborating systems' improvement strategies by interviewing six to ten key system leaders in each system and by collecting system-level planning and implementation documents in October of each year. The leaders were from a number of system units that had a stake in mathematics teaching and learning. They included Curriculum and Instruction that is responsible for selecting instructional materials and for providing professional development for teacher and coaches, Leadership that that is responsible for providing professional development for school leaders and for holding school leaders accountable, ELL that is responsible for supporting the learning of English Language Learners, Special Education that is responsible for supporting the learning of students who receive special education services, and Research and Evaluation that is responsible for generating and analyzing data on students, teachers, schools, and the district.

In addition to asking about current initiatives in middle-grades mathematics, it proved useful to include interview questions that focused on student demographics, the impact of regional and national policies, and the historical context of the system including prior reform initiatives and previous mathematics instructional materials and assessments. (Interview protocols are downloadable at http://vanderbi.lt/mist).

The transcribed interviews and the artifacts can be analyzed through an inductive coding process in order to discern broad consistencies across participants in each system. The goal in conducting these analyses is to clarify the intended or envisioned practices of members of particular role groups (e.g., teachers, coaches, principals), the intended means of supporting the learning of members of

those groups, and system leaders' rationales for why the supports might enable members of each role group to develop the envisioned forms of practice.

In the MIST study, we reported our findings for each collaborating system in a five-page document. This *System Design Document* named each district strategy and described the intended supports and accountability relations for members of each role group. We shared this document with system leaders to determine whether it accurately represented their plan for instructional improvement. We made revisions until the district leaders agreed that the document accurately represented their intended plan.

System Design Documents serve four useful purposes. First, they are useful in preparing for the next phase of a cycle that involves collecting data to learn how each system's intended strategies are being implemented in schools. Second, the major strategies identified in each document provide a framework for organizing the feedback given to the system leaders about how their improvement strategies are playing out. Third, system leaders who participated in the MIST study reported that found these documents useful in clarifying their improvement strategies with others across the system. Finally, the *System Design Documents* produced in successive cycles provide a record of changes in a system's improvement policies over time, thus enabling the system leaders to monitor progress and researchers to document the influence of their recommendations on the improvement strategies that system leaders attempted to implement in the next cycle.

To illustrate, we refer to the *System Design Document* we created for District B, one of the four participating districts, during our first year of working with the district. (Table 1 provides a summary of District B's *System Design Document*, 2007-2008). The overall goal of the instructional improvement effort in District B was to ensure that all students had opportunities to learn through engagement with a rigorous mathematics curriculum, that teachers and school leaders had high expectations for students' learning, and that achievement disparities between White students and traditionally underserved groups of students were eliminated. District B was in its first year of implementing an inquiry-oriented mathematics curriculum. To support this implementation, the district had assigned a mathematics coach to each middle school the previous year and had provided them with a significant amount of professional development that focused on both teaching the new curriculum effectively and coaching other mathematics teachers at their schools. Each coach taught for half of the school day and served as a coach for the remainder of the day.

==========INSERT TABLE 1 ABOUT HERE ======================

The first improvement strategy that we identified was to support principals' and mathematics coaches' development as instructional leaders who worked together to improve the quality of mathematics instruction. Principals were expected to observe classroom instructional regularly to assess the quality of teachers' instructional practices and determine their needs based on these observations.

Principals received professional development on observing and assessing the quality of mathematics instruction, and were expected to meet with the mathematics coach at their school every week to discuss the quality of classroom instruction and assess teachers' needs.

The second strategy was to support teachers in teaching the inquiry-oriented curriculum effectively. Supports for teachers' learning included teacher professional development provided by the mathematics coaches and a district Curriculum Framework that aligned the curriculum with the state standards and provided guidance on differentiating instruction for particular groups of students, especially English Language Learners and special education students.

We used the Interpretive Framework described above to assess the strengths and limitations of these two improvement strategies. District leaders clearly and consistently articulated the forms of practice they intended teachers, coaches, and principals would develop (e.g. principals were to observe classrooms and provide feedback to improve instructional practices). In addition, these intended forms of practice were compatible with the district's overall goal of supporting teachers' development of ambitious instructional practices. However, we considered it unlikely that the supports for various role groups' learning would be adequate.

With regard to the first strategy, principals would have to distinguish between weak and strong enactments of ambitious instructional practices if they were to give teachers effective feedback. The supports for principals' learning included professional development on observing classroom instruction. We questioned whether these ongoing intentional learning events would be effective because they focused on characteristics of high quality instruction that were independent of subject matter area, and because these characteristics were relatively global. Principals were also expected to meet regularly with the mathematics coach to discuss the quality of classroom instruction. Although these discussions might focus on content-specific instructional practices, we doubted whether the resulting incidental learning opportunities would be adequate. In addition, the coaches were new to the role and it was not clear that they had developed sufficient expertise to support principals in assessing the quality of instruction.

With regard to the second strategy, the effective implementation of the inquiry-oriented curriculum that the district had adopted required that most teachers significantly reorganize their instructional practices. Teachers participated in *ongoing intentional learning events*, four days of district professional development led by the mathematics coaches. However, it was not clear that mathematics coaches had developed the expertise to lead this professional development effectively given that they were also teaching the new curriculum for the first time.

B) Documenting How Instructional Improvement Strategies are Being Implemented

The next phase of the design cycle involves collecting data to document how each system's strategies are playing out in schools and classrooms. In the MIST study, we collected multiple types of data to document the four systems' instructional improvement efforts: audio-recorded interviews conducted with the 200 participants; on-line surveys for teachers, coaches, and school leaders; video-recordings of two consecutive lessons in the 120 participating teachers' classrooms, coded using the Instructional Quality Assessment (IQA) (Boston, 2012; Matsumura, Garnier, Slater, & Boston, 2008); teachers' and coaches' scores on the Mathematics Knowledge for Teaching (MKT) instrument (Hill, Schilling, & Ball, 2004); video-recordings of select district professional development; audio-recordings of teacher collaborative planning meetings; and an on-line assessment of teacher networks completed by all mathematics teachers in the participating schools. In addition, the districts provided us with access to mathematics achievement data for students in the participating 120 teachers' classrooms. The interviews and online surveys focused on the school and district settings in which the participating teachers and school leaders worked and gave particular attention to the formal and informal supports on which they could draw to improve their practices, as well as to whom they were accountable and for what they were accountable.

As we had only three months to analyze data before district leaders began planning strategies for the following school year, we limited the data we analyzed to provide feedback about how districts' strategies were being implemented to the audio-recorded interviews conducted with the 50 participants in each district. (As our collaboration with each district continued over four years, we were able to share additional findings from other data sources, for example video-recordings of classroom instruction, in subsequent reports as they became available.)

One of the challenges when conducting a system-level design study is to analyze a large amount of data in a relatively short period of time while ensuring that the findings shared with system leaders are reliable. In this context, an important criterion for reliability is that claims about how improvement strategies are being implemented can be justified by backtracking through successive steps of the analysis to the raw data. This method involves using a series of structured tools to first summarize transcriptions of each participant interview, and then to triangulate and synthesize the responses both across participants in each school and across teachers, coaches, and school leaders in each collaborating system.

In MIST, a team member completed an Interview Summary Form (ISF) for each interview (teacher, coach, school leader, system leader). The ISF summarized each participant's response to interview questions that were central to understanding how improvement strategies were playing out in schools. This information was then synthesized across all participants in a school using the School Summary Form (SSF). This required the triangulation of participant responses at each school, citing evidence from the ISFs. Additional forms included a Principal Summary Form (PSF), a Coach Summary Form (CSF) and a Teacher Summary Form (TSF) that were used to synthesize information across members of a role

group in a system (i.e., the TSF synthesized the interview summary forms for all the participating teachers in a system).

Once this initial analysis was complete, we returned to the *System Design Document*, and we identified gaps between each system's intended strategies and the strategies as they were being implemented in schools. We then examined why strategies were playing out as we had documented rather than as intended by focusing on the actual learning opportunities and press for improvement for members of each role group. In developing these explanations, we used one of our conceptual tools, the interpretive framework which differentiates between four general types of supports.

The final step in the analysis involved developing recommendations for how system leaders might revise their improvement strategies to make them more effective. In doing so, we drew on the conjectures about supports and accountability relations that comprised the current iteration of our theory of action for instructional improvement. It also proved essential to take account of each collaborating system's current capacity to support the learning of members of particular role groups. The resulting recommendations proposed feasible strategies for supporting teachers', coaches,' and school leaders' improvement of their practices.

As an illustration, our analysis of data collected during our first year of collaborating with District B indicated that the knowledge and instructional practices of the school-based mathematics coaches were only slightly more advanced than those of the teachers they were expected to support even though they had received extensive professional development. In addition, few if any mathematics teachers had developed relatively accomplished instructional practices. Further, the district had only three district-level mathematics specialists (members of Curriculum and Instruction) who were expected to fulfill several different roles and responsibilities while serving 32 middle-grades schools. This lack of instructional expertise was a major constraint that we had to take account of when making recommendations about supports for teachers' and for coaches' learning. One of our recommendations therefore included leveraging the expertise of the three mathematics specialists by making their work in supporting the coaches' learning a priority.

C) Sharing Findings and Recommendations With System Leaders

We have emphasized that a system-level design study involves a genuine partnership with system leaders in which the leaders have the ultimate authority for making decisions about improvement strategies. It is therefore important for the researchers to develop a method for sharing findings and recommendations that is both feasible and relevant to system leaders' current concerns. In the MIST study, a two-step process for communicating findings and negotiating future improvement strategies proved to be relatively effective.

The first step involved preparing a *System Feedback and Recommendations Report* of approximately 15 single-spaced pages for the leaders of each collaborating system. These reports built directly on the *System Design Documents* and were intentionally structured around the district's major strategies so that they related directly to the work district leaders were attempting to accomplish. For each strategy reported in the *System Design Document*, we reiterated the envisioned forms of practice that constituted the goal of the strategy and described the intended supports and accountability relations for the development of the envisioned practices. We then reported our findings about how that strategy was playing out in schools, explained why this was the case, and made our recommendations for adjusting the strategy. Based on our experience in the MIST study, we believe that this way of organizing reports for system leaders provides a useful model for others conducting system-level design studies. Redacted versions of reports produced in the MIST study are available at the MIST website (http://vanderbi.lt/mist).

The second step in sharing findings and recommendations with system leaders was a two-hour meeting with the leaders of each system scheduled approximately one week after we sent them the *System Feedback and Recommendations Report*. The intent of these meetings was to clarify the findings and to have a genuine conversation about their implications for the system's improvement strategies. We therefore recommend that researchers explicitly negotiate norms for these meetings with system leaders, and that they speak from notes rather than PowerPoint in order to encourage an open discussion.

In the case of MIST, these meetings usually included the head of Curriculum and Instruction responsible for all content areas, the head of the Mathematics Department, the district mathematics specialists (who work with the mathematics coaches and support schools), the head of Leadership, and leadership specialists who support and assess school leaders. In one district the superintendent attended the feedback sessions. These meetings were typically very productive. In every instance, the conversation was an open dialogue about the current status of the district's improvement efforts and about possible adjustments to those efforts.

D) Assessing the Influence of Recommendations on Collaborating System's Instructional Improvement Strategies

The first phase of the next data collection, analysis, and feedback cycle involves interviewing system leaders again to document their revised instructional improvement strategies. The influence of recommendations made to system leaders can be assessed by comparing their revised and prior improvement strategies. As we have noted, assessing the influence of the recommendations is important both because the pragmatic goal of a system-level design is to contribute to the collaborating systems' instructional improvement efforts, and because researchers

have opportunities to test the conjectures that comprise their theory of action when system leaders act on their recommendations. A priori, sharing a written report and conducting a single meeting with system leaders to discuss its implications might appear to be a relatively weak mechanism for influencing system-wide improvement strategies. We were therefore gratified to find that the leaders in all four systems that participated in the MIST study revised their improvement strategies based on many of our recommendations. The data we collected the following year documented how the policy revisions we recommended were actually playing out in school and classrooms. We could therefore use these data to test, revise and thus improve our conjectures about the learning of members of different role groups and the means of supporting their learning. Thus, the collaborative partnerships in which we became co-designers of district improvement policies with district leaders enabled us to enact iterative cycles of design and analysis that are characteristic of the design research methodology.

There are several reasons why we believe this limited collaboration proved to be sufficient. First, we selected districts whose goals for students' mathematical learning and for teachers' improvement of their instructional practices were broadly compatible with those that we intended to investigate. Second, we prioritized the development and maintenance of relationships of trust with district leaders and school personnel. Thus, during the first year of the collaboration, we strove to produce feedback reports that district and schools leaders would view as extremely relevant and useful to their work. It was because districts leaders found this and subsequent reports useful that they were willing to continue to spend time with us three times a year (fall interview, January interview, and May feedback session) and to allocate resources to assist our data collection efforts. This in turn enabled us to achieve almost 100% success in all aspect of our data collection each year in all four districts.

This approach of developing, testing, and refining theory by conducting tightly integrated cycles of analysis and (policy) design is at the heart of the design research methodology (Cobb et al., 2003). On the one hand, we revised and elaborated the conjectures that comprised our evolving theory of action in the course of the analysis and feedback process. On the other hand, our evolving conjectures informed the formulation of the specific feedback recommendations we made to the districts. In a very real sense, the design for system-wide instructional improvement that is implemented was co-constructed by system leaders and the researchers. It is useful to distinguish between the co-constructed designs that are specific to a particular system and researchers' conjectures about the process of supporting system-wide instructional improvement more generally. On the one hand, these latter conjectures comprise a theory of action that can be used to make recommendations to system leaders. On the other hand, occasions when system leaders act on these recommendations constitute opportunities for the researchers to test and revise their conjectures and thus contribute to the development of a generalizable theory of action for system-wide instructional improvement in mathematics.

Testing and Revising Conjectures that Comprise a Theory of Action for System-Wide Instructional Improvement

To this point, we have focused on the pragmatic objective of providing leaders of the collaborating systems with timely feedback about how their improvement strategies are actually playing out that can inform the revision of their instructional improvement efforts. We now consider the theoretical objective of contributing to a generalizable theory of action for system-wide instructional improvement in mathematics. In doing so, we draw on our experience in the MIST study by discussing three types of evidence that can inform the revision of conjectures that comprise the theory of action: findings from feedback analyses about how the collaborating systems' instructional improvement strategies are being implemented, the current research literature, and the findings of retrospective analyses conducted by drawing on the multiple sources of data collected in each cycle.

Findings About the Districts' Instructional Improvement Strategies

When researchers formulate recommendations to collaborating educational systems, they necessarily have to address concrete organizational design challenges by proposing how the systems might support and hold members of particular role groups accountable for improving their practices. Addressing these challenges is a primary context for researchers' learning. Furthermore, researcher can step back after completing each data collection, analysis, and feedback cycle and frame the findings and recommendations for the collaborating systems as cases of attempting to support instructional improvement at scale. In doing so, it is important to determine whether any of the recommendations to a particular system represent refinements or elaborations of current conjectures, and if they do whether they might have more general implications and under what conditions. For example, the constraint of limited instructional expertise that we identified in District B proved to be a constraint in two of the other three collaborating districts. The recommendations we made to these districts for supporting teachers' and coaches' learning could therefore inform the revision of our initial conjectures for instructional improvement in districts that are constrained by limited instructional expertise.

Research Literature

As we have noted, relevant research that can inform the design of instructional improvement strategies becomes increasingly thin the further one moves away from the classroom (Cobb et al., in press; Honig, 2012). Nonetheless, findings reported in the literature can, on occasion, provided evidence for the revision

of current conjectures. This possibility is becoming increasingly likely as system-level design studies and closely related approaches become more common.

Retrospective Analyses

The retrospective analysis of data collected during successive design and analysis cycles is a key aspect of design studies conducted at any level. In the case of system-level design studies, a primary goal of retrospective analyses is to investigate key conjectures of the theory of action for instructional improvement. Based on our work in the MIST study, we recommend that mutually informing lines of retrospective analyses be established that focus on the major types of supports conjectured to be important for instructional improvement (e.g., teacher collaborative time, teacher networks, mathematics coaching, school instructional leadership).

As we have indicated, the types of data that can be analyze to give collaborating systems feedback about how their improvement strategies are playing out is constrained by the need to ensure that the feedback is timely and can inform system leaders' revision of their strategies. Retrospective analyses that can inform the revision of the theory of action draw on a range of additional types of data are collected during each data collection, analysis, and feedback cycle. The primary concern when making decisions about the types of data to collect is that the key constructs of each conjecture are assessed including the relevant aspects of teachers' knowledge and instructional practices. For example, if the vision of high-quality mathematics instruction that constitutes the goal for teachers' learning requires that teachers deepen their mathematical knowledge, then it is important to include an appropriate measure of this knowledge. Similarly, if teachers' informal professional networks are conjectured to be an important support for their learning, then it is important to develop instruments for assessing the relevant aspects of their networks (e.g., who teachers turn to for instructional advice, frequency of their interactions with those people, and content of their interactions).

The MIST team is currently conducting five interrelated lines of analysis that focus on district-level and school-level teacher professional development (including mathematics teacher collaborative meetings), teacher networks, mathematics coaching, school instructional leadership, and district instructional leadership. We discuss the current version of our theory of action for instructional improvement in mathematics in the next section of this chapter.

MIST's Current Theory of Action for Instructional Improvement in Middle-Grades Mathematics

Presenting the current iteration of our theory of action in any detail is beyond the scope of this chapter and we refer the reader to Cobb and Jackson (2011).

To illustrate our current conjectures, we focus on one component of the theory of action, school instructional leadership.

Our initial conjectures about school instructional leadership were relatively global and did not differentiate between the practices of mathematics coaches and school leaders. These conjectures indicated the importance of school leaders developing relatively sophisticated visions of high-quality mathematics instruction and both supporting and holding mathematics teachers accountable for developing high-quality instructional practices. Our revised conjectures indicate the potential value of a distributed model of school instructional leadership in which coaches and district mathematics specialists are primarily responsible for supporting teachers' learning, and school leaders are responsible for pressing and holding teachers accountable for developing the intended instructional practices (Elmore, 2006; Spillane, Halverson, & Diamond, 2004). In addition, our current conjectures specify three leadership practices that might be feasible goals for school leaders' learning. Two of these practices—observing mathematics instruction and providing feedback, and participating in mathematics teacher collaborative meetings—aim at pressing teachers to develop the intended forms of practice and providing teachers with adequate support. The third practice concerns the development of productive relationships with coaches.

We conjecture that by observing instruction and providing teachers with informed feedback, school leaders can both communicate expectations and hold teachers accountable for improving classroom instruction. We also conjecture that it is important that the feedback be specific to the instructional practices on which school and district teacher professional development focuses. However, the extent to which school leaders' feedback accomplishes these goals depends crucially on the professional development they receive.

We conjecture that school leaders' participation in mathematics teacher collaborative meetings signals the importance of teacher collaboration, enables school leaders to hold teachers accountable for using collaborative time productively, and constitutes a context for school leaders' learning, thus better positioning them to give productive feedback after observing instruction and to procure appropriate resources for teachers. In this regard, a meta-analysis conducted by Robinson, Lloyd, and Rowe (2008) found that school leaders' participation in teacher professional development is strongly associated with improvements in student achievement.

Findings of a retrospective analysis indicate that coaches' effectiveness in supporting teachers' learning depends on school leaders assuming shared responsibility for instructional improvement with them (Gibbons, Garrison, & Cobb, 2010). We therefore conjecture that it is important that school leaders understand the district-wide goals for students' mathematical learning and the guiding vision of high-quality instruction, and that they appreciate the critical role of coaches in supporting teachers' learning. In the course of our collaboration with the districts, we have documented several cases in which principals have assigned additional duties to coaches that took them away from their work with teachers (e.g., analyz-

ing data to identify struggling students, tutoring struggling students). Our observations also indicate that principals protect coaches' time when they understand the coaches' role in the improvement effort. We conjecture that the development of shared responsibility for instructional improvement is facilitated if school leaders and coaches meet regularly to share their observations about the quality of teachers' instructional practices, discuss how the coach's work with teachers is progressing, jointly select teachers with whom the coach should work, and plan for future work with groups of teachers.

The ongoing analyses we have conducted while developing feedback for the collaborating districts indicate that it is challenging for school leaders, most of whom are not mathematics specialists, to develop the three instructional leadership practices that we have described. As a consequence, we have also developed conjectures about the nature of professional development that might support their development of these practices.

First, we conjecture that if school leaders are to effectively and realistically press teachers to improve the quality of instruction, professional development for school leaders should enable them to recognize the instructional practices that are the focus of teacher professional development, and to distinguish between low- and high-quality enactments of those practices. We also conjecture that a consistent emphasis on the same instructional practices across teacher, coach, and school leader professional development will contribute to the development of compatible visions of high-quality instruction and to the alignment of supports for teachers' learning.

Second, we conjecture that professional development should attend explicitly to how to provide feedback to teachers that communicates expectations for ambitious instruction. This might involve school leaders and district mathematics specialists observing instruction or watching video-recordings of specific phases of lessons and discussing the feedback they would provide.

Third, we conjecture that professional development should clarify the role of coaches and mathematics teacher collaborative meetings in supporting teachers' development of ambitious instructional practices. We have documented several cases in which a school leader has taken over the agenda mathematics teacher meetings to the detriment of the participating teachers' learning. We therefore conjecture that it is important to give particular attention to how the distribution of instructional leadership between coaches and school leaders should reflect their complementary areas of expertise (Elmore, 2006).

The contrast between our initial and current conjectures for school leadership is representative of the changes we have made as we have revised and elaborated our initial conjectures. The level of specificity of our current conjectures is essential if we are to provide district leaders with *actionable* guidance on how they might support instructional improvement in mathematics on a large scale. We regard the current iteration of our theory of action as a work in progress and are further testing and revising our conjectures as we continue to collaborate with two of the four districts for a further four years.

Conclusion

Our purpose in this chapter has been to describe a design research approach for studying and supporting improvements in the quality of mathematics teaching on a large scale. The aim of this methodology is to both provide the leaders of educational systems, such as urban school districts in the U.S., with feedback that can inform their instructional improvement efforts and to contribute to the development of a generalizable theory of action for large-scale instructional improvement in mathematics. The successful use of the methodology depends crucially on researchers establishing a genuine collaborative partnership with educational leaders such that researchers become co-designers of instructional improvement policies. Only then is it possible for researchers to test and revise their conjectures about supports for instructional improvement by conducting successive data collection, analysis, and feedback cycles.

We noted early in this chapter that research in mathematics education has made considerable progress in recent years, but that the findings of this work have had little impact on the quality of mathematics instruction and thus student learning in most classrooms. Design studies of the type that we have described and illustrated are clearly non-trivial undertakings. The value of this methodology derives from the way in which it enables us to test, revise, and thus improve our understanding of what it takes to support large-scale instructional improvement in mathematics..

References

Argyris, C., & Schon, D.A. (1978). *Organizational learning: A theory of action perspective.* Reading, MA: Addison-Wesley.

Argyris, C., & Schon, D. (1974). *Theory of practice*. San Francisco: Jossey-Bass.

Blumenfeld, P., Fishman, B. J., Krajcik, J. S., Marx, R., & Soloway, E. (2000). Creating usable innovations in systemic reform: Scaling-up technology - embedded project-based science in urban schools. *Educational Psychologist, 35*, 149-164.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3-15.

Boston, M. D. (2012). Assessing the quality of mathematics instruction. *Elementary School Journal, 113*(1), 76-104.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*, 141-178.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 32-42.

Bruner, J. (1987). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Bryk, A. S. (2009). Support a science of performance improvement. *Phi Delta Kappan*, 90(8), 597-600.

Bryk, A. S., & Gomez, L. M. (2008). Reinventing a research and development capacity. In F. Hess (Ed.), *The future of educational entrepreneurship: Possibilities for school reform* (pp. 181-187). Cambridge, MA: Harvard Education Press.

Cobb, P. (2000). Conducting teaching experiments in collaboration with teachers. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 307-334). Mahwah, NJ: Erlbaum.

Cobb, P. (2002). Reasoning with tools and inscriptions. *Journal of the Learning Sciences, 11*, 187-216.

Cobb, P, Confrey, J., diSessa, A., Leher, R., Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9-13.

Cobb, P., & Jackson, K. (2011). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale. *Mathematics Teacher Education and Development, 13*(1), 6-33.

Cobb, P., & Jackson, K. (2012). Analyzing educational policies: A learning design perspective. *The Journal of the Learning Sciences*, 21(4), 487-521.

Cobb, P., Jackson, K., Smith, T., Sorum, M., & Henrick, E. (in press). Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale. In B. J. Fishman, W. R. Penuel, A.-R. Allen & B. H. Cheng (Eds.), *Design based implementation research. National Society for the Study of Education Yearbook.*

Cobb, P., & McClain, K. (2001). An approach for supporting teachers' learning in social context. In F.-L. Lin & T. Cooney (Eds.), *Making sense of mathematics teacher education* (pp. 207-232). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cobb, P., McClain, K., Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and school district. *Educational Researcher, 32*(6), 13-24.

Cobb, P., & Smith, T. (2008). The challenge of scale: Designing schools and districts as learning organizations for instructional improvement in mathematics. In T. Wood, B. Jaworski, K. Krainer, P. Sullivan & D. Tirosh (Eds.), *International handbook of mathematics teacher education* (Vol. 3, pp. 231-254). Rotterdam, Netherlands

Cobb, P., & Steffe, L. P. (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education, 14*, 83–94.

Cobb, P., Zhao, Q., & Dean, C. (2009). Conducting design experiments to support teachers' learning: A reflection from the field. *Journal of the Learning Sciences, 18*, 165-199.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3-12.

Coburn, C. E., & Russell, J. L. (2008). District policy and teachers' social networks. *Educational Evaluation and Policy Analysis, 30*(3), 203-235.

Coburn, C. E., & Stein, M. K. (Eds.). (2010). *Research and practice in education: Building alliances, bridging the divide*. New York: Rowman & Littlefield Publishing Group.

Cohen, D. K., & Barnes, C. A. (1993). Pedagogy and policy. In D. K. Cohen, M. W. McLaughlin & J. E. Talbert (Eds.), *Teaching for understanding: Challenges for policy and practice* (pp. 207-239). San Francisco: Jossey Bass.

Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, 102, 294-343.

Cole, M. (1996). *Cultural psychology*. Cambridge, MA: Belknap Press of Harvard University Press.

Collins, A. (1992). Toward a design science of education. In E. Scanon & T. O'Shey (Eds.), *New directions in educational technology* (pp. 15-22). New York: Springer.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13*, 15-42.

Common Core State Standards Initiative (2010). *Common Core State Standards for Mathematics*. Retrieved from http://www.corestandards.org

Confrey, J., & Lachance, A. (2000). Transformative teaching experiments through conjecture-driven research design. In E. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 231–266). Mahwah, NJ: Lawrence Erlbaum Associates.

Crockett, M. D. (2007). Teacher professional development as a critical resource in school reform. *Journal of Curriculum Studies*, 39, 253-263.

Darling-Hammond, L. (2007). The flat earth and education: How America's commitment to equity will determine our future. *Educational Researcher, 36*(6), 318-334.

Design-Based Research Collaborative (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5-8.

Desimone, L., Porter, A. C., Garet, M., Suk Yoon, K., & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year study. *Educational Evaluation and Policy Analysis*, 24, 81-112.

Easton, J. (2010, June). *New research initiatives for IES.* Keynote Address at the IES Research Conference, National Harbor, MD.

Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences, 11*, 105-121.

Elmore, R. F. (1979-80). Backward mapping: Implementation research and policy decisions. *Political Science Quarterly*, 94, 601-616.

Elmore, R. F. (2004). School reform from the inside out. Cambridge, MA: Harvard Education Press.

Elmore, R. F. (2006). Leadership as the practice of improvement, *OECD International Conference on Perspectives on Leadership for Systemic Improvement*. London.

Feldman, M. S. (2000). Organizational routines as a source of continuous change. *Organization Science, 11*, 611-629.

Feldman, M. S. (2004). Resources in emerging structures and processes of change. *Organization Science, 15*, 295-309.

Feldman, M. S., & Pentland, B. T. (2003). *Reconceptualizing organizational routines as a source of flexibility and change.* Administrative Science Quarterly, 48(1), 1-19.

Fink, E., & Resnick, L.B. (2001). Developing principals as instructional leaders. *Phi Delta Kappan, 82*, 578-606.

Fishman, B., Marx, R. W., Blumenfeld, P., & Krajcik, J. S. (2004). Creating a framework for research on systemic technology innovations. *Journal of the Learning Sciences, 13*, 43-76.

Forman, E. A. (2003). A sociocultural approach to mathematics reform: Speaking, inscribing, and doing mathematics within communities of practice. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 333-352). Reston, VA: National Council of Teachers of Mathematics.

Franke, M. L., & Kazemi, E. (2001). Learning to teach mathematics: Developing a focus on students' mathematical thinking. *Theory into Practice*, 40, 102-109.

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225-256). Greenwich, CT: Information Age Publishers.

Gamoran, A., Anderson, C. W., Quiroz, P. A., Secada, W. G., Williams, T., & Ashmann, S. (2003). *Transforming teaching in math and science: How schools and districts can support change*. New York: Teachers College Press.

Gamoran, A., Secada, W. G., & Marrett, C. B. (2000). The organizational context of teaching and learning: Changing theoretical perspectives. In M. T. Hallinan (Ed.), *Handbook of sociology of education* (pp. 37-63). New York: Kluwer Academic/Plenum Publishers.

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, 915-945.

Gibbons, L. K., Garrison, A. L., & Cobb, P. A. (2010, October). *Teacher networks and the role of the mathematics coach: how institutional factors influence coach centrality*. Paper presented at the annual conference of the North American Chapter of the International Group for the Psychology of Mathematics Education, Columbus, OH.

Gravemeijer, K. (1994). Educational development and developmental research. *Journal for Research in Mathematics Education, 25*, 443-471.

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184-205.

Heilig, J. V., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis, 30*(2), 75-110.

Hess, Frederick, M. (1999). *Spinning wheels: The politics of urban school reform*. Washington DC: Brookings Institution.

Hill, H.C., Schilling, S.G., & Ball, D.L. (2004) Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, *105*(1), 11-30.

Honig, M. I. (2012). District central office leadership as teaching: How central office administrators support principals' development as instructional leaders. *Educational Administration Quarterly*, 48, 733-774.

Horn, I. S. (2005). Learning on the job: A situated account of teacher learning in high school mathematics departments. *Cognition and Instruction*, 23, 207-236.

Horn, I. S. (2012). *Strength in numbers: Collaborative learning in secondary mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (in press). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education.*

Kazemi, E., & Franke, M. L. (2004). Teacher learning in mathematics: Using student work to promote collective inquiry. *Journal of Mathematics Teacher Education*, 7, 203-225.

Kazemi, E., Franke, M. L., & Lampert, M. (2009). Developing pedagogies in teacher education to support novice teachers' ability to enact ambitious instruction. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 12-30). Palmerston North, NZ: MERGA.

Kazemi, E., & Hubbard, A. (2008). New directions for the design and study of professional development: Attending to the coevolution of teachers' participation across contexts*. Journal of Teacher Education, 59, 428-441.*

Kilpatrick, J., Martin, W. G., & Schifter, D. (Eds.). (2003). *A research companion to principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Kozulin, A. (1990). *Vygotsky's psychology: A biography of ideas*. Cambridge: Harvard University Press.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M. L. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 129-141). New York: Springer.

Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *The Elementary School Journal,* 109(5), 491-509.

Lampert, M., & Graziani, F. (2010). Instructional activities as a tool for teachers' and teacher educators' learning. *Elementary School Journal*, *109*, 491–509.

Lave, J. (1993). The practice of learning. In S. Chaiklin & J. Lave (Eds.), Understanding practice:
Perspectives on activity and context (pp. 3–32). Cambridge, England: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* London: Cambridge University Press.

Lehrer, R., & Lesh, R. (2003). Mathematical learning. In W. Reynolds & G. Miller (Eds.), *Comprehensive handbook of psychology* (Vol. 7, pp. 357-391). New York: John Wiley.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal*, 41, 635-679.

Lehrer, R., Schauble, L., & Penner, D. (2000). The inter-related development of inscriptions and conceptual understanding. In P. Cobb, E. Yackel & K. McClain (Eds.), *Symbolizing, mathematizing, and communicating: Perspectives on discourse, tools, and instructional design* (pp. 325-360). Mahwah, NJ: Erlbaum.

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*, 917-946.

Lobato, J. (2003). How design experiments can inform a rethinking of transfer and vice versa. *Educational Researcher, 32*(1), 17-20.

Mangin, M. M. (2007). Facilitating elementary principals' support for instructional teacher leadership. *Educational Administration Quarterly*, 43, 319-357.

Marks, H. M., & Louis, K. S. (1997). Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic performance. *Educational Evaluation and Policy Analysis, 19*, 245-275.

Matsumura, L. C., Garnier, H., Slater, S. C., & Boston, M. D. (2008). Toward measuring instructional interactions 'at-scale'. *Educational Assessment*, 13(4), 267-300.

McClain, K. (2002). Teacher's and students' understanding: The role of tool use in communication. *Journal of the Learning Sciences, 11*, 217-249.

McLaughlin, M. W., & Mitra, D. (2004, April). *The cycle of inquiry as the engine of school reform: Lessons from the Bay Area School Reform Collaborative.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education, 29*, 121-142.

National Council of Teachers of Mathematics. (1989). *The curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Nelson, B. S., & Sassi, A. (2005). *The effective principal: Instructional leadership for high-quality learning.* New York: Teachers College Press.

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions* (pp. 47-87). New York: Cambridge University Press.

Penuel, W. R., Frank, K. A., & Krause, A. (2006). *The distribution of resources and expertise and the implementation of schoolwide reform initiatives*. Paper presented at the Seventh International Conference of the Learning Sciences, Bloomington, IN.

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, 40(7), 331-337.

Remillard, J. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75, 211-246.

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*, 635-674.

Roderick, M., Easton, J. Q., & Sebring, P. B. (2009). *The Consortium on Chicago School Research: A New Model for the Role of Research in Supporting Urban School Reform*. Chicago: The Consortium on Chicago School Research at the University of Chicago Urban Education Institute.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* Oxford: Oxford University Press.

Rogoff, B. (1997). Evaluating development in the process of participation: Theory, methods, and practice building on each other. In E. Amsel & A. Renninger (Eds.), *Change and development: Issues of theory, application, and method* (pp. 265-285). Hillsdale, NJ: Erlbaum.

Schoenfeld, A. H. (2006). Design experiments. In J. L. Green, G. Camilli, P. B. Ellmore, & A. Skukauskaite (Eds.), *Handbook of Complementary Methods in Education Research* (pp. 193-206). Washington, DC: American Educational Research Association.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. New York: Cambridge University Press.

Sherer, J. Z., & Spillane, J. P. (2011). Constancy and change in work practice in schools: The role of organizational routines. *Teachers College Record, 113*(3), 611-657.

Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Curriculum Studies, 31*, 143-175.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2004). Distributed leadership: Towards a theory of school leadership practice. *Journal of Curriculum Studies, 36*, 3-34.

Spillane, J. P., Mesler, L., Croegaert, C., & Sherer, J. Z. (2007, August). Coupling administrative practice with the technical core and external regulation: The role of organizational routines. Paper presented at the annual meeting of the European Association for Research on Learning and Instruction, Budapest, Hungary.

Spillane, J. P., Reiser, B., & Gomez, L. M. (2006). Policy implementation and cognition: The role of human, social, and distributed cognition in framing policy implementation. In M. I. Honig (Ed.), *New directions in educational policy implementation* (pp. 47-64). Albany, NY: State University of New York Press.

Spillane, J. P., & Thompson, C. L. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis, 19*, 185-203.

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction*, 25(2), 161-217.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 267-307). Mahwah, NJ: Erlbaum.

Stein, M. K. (2004). Studying the influence and impact of standards: The role of districts in teacher capacity. In J. Ferrini-Mundy & F. K. Lester, Jr. (Eds.), *Proceedings of the National Council of Teachers of Mathematics Research Catalyst Conference* (pp. 83-98). Reston, VA: National Council of Teachers of Mathematics.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313-340.

Stein, M. K., & Nelson, B. S. (2003). Leadership content knowledge. *Educational Evaluation and Policy Analysis, 25*(4), 423-448.

Stein, M. K., Silver, E. A., & Smith, M. S. (1998). Mathematics reform and teacher development: A community of practice perspective. In J. G. Greeno & S. V. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 17-52). Mahwah, NJ: Lawerence Erlbaum Associates.

Stephan, M., Bowers, J., & Cobb, P. (Eds.). (2003). Supporting students' development of measuring conceptions: Analyzing students' learning in social context. Journal *for Research in Mathematics Education Monograph No. 12*. Reston, VA: National Council of Teachers of Mathematics.

Supovitz, J. A. (2006). *The case for district-based reform*. Cambridge, MA: Harvard University Press.

Tyack, D., & Tobin, W. (1995). The "Grammar" of schooling: Why has it been so hard to change? *American Educational Research Journal, 31*, 453-479.

U.S. Department of Education. (2008). *The final report of the National Mathematics Advisory Panel.* Washington, DC: Author.

Wenger, E. (1998). Communities of practice. New York: Cambridge University Press.

Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Washington, DC: Association for Curriculum and Supervision.

van der Veer, R., & Valsiner, J. (1991). Understanding Vygotsky: A quest for synthesis. Cambridge, MA: Blackwell.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Table 1. Summary of a System Design Document for District B, 2007-2008 School Year

District B Instructional Improvement Goals

Ensure that all students have opportunities to learn through engagement with a rigorous curriculum, that teachers and school leaders have high expectations for students' learning, and that achievement gaps between White students and traditionally underserved groups of students are eliminated.

Improvement Strategies	Supports for Role Groups to Develop the Intended Forms of Practice
1) Develop principals and coaches who work together to improve instruction	- Professional development for principals on observing classroom and providing feedback to teachers
	- Principal and the math coach are required to meet weekly to discuss classroom instruction and supports for teachers
	- Professional development for math coaches
2) Support teachers in teaching a rigorous mathematics curriculum effectively	- Professional development for teachers on the inquiry-oriented curriculum
	- A comprehensive curriculum framework to support the implementation of the rigorous curriculum