Investigating What It Takes to Improve the Quality of Mathematics Teaching and Learning on a Large Scale[1]

Paul Cobb and the MIST Team

Research on the teaching and learning of mathematics has made significant progress in recent years. However, this work has had only limited impact on classroom instruction in many countries including the US. I report on an investigation in which we collaborated with mathematics teachers, school leaders, and the leaders in several large urban school systems in the US for eight years to investigate what it takes to support improvements in the quality of instruction and thus students' learning on a large scale. Our findings from this work take the form of an empirically-grounded theory of action for instructional improvement at scale that spans from the classroom to system instructional leadership and encompasses: curriculum materials and assessments; pull-out teacher professional development; school-based teacher collaborative meetings; coaches' practices in providing job-embedded support for teachers' learning; school leaders' practices as instructional leaders in mathematics; and system leaders' practices in supporting the development of school-level capacity for instructional improvement. In outlining the key aspects of the theory of actions, I also discuss smaller, embedded studies that we conducted to investigate specific conjectures about supporting professional leaders' development of more effective practices.

Introduction

For the last 15 years, my colleagues and I have focused on the question of what it takes to improve the quality of mathematics teaching on a large scale. We came to see that addressing this question involves tackling two closely related issues. The first is to identify a coherent set of potentially productive instructional improvement strategies, and the second is to learn how to support the effective implementation of those strategies across a range of school and school system contexts. In this paper, I give an overview of our findings as they relate to the first of these issues. We began to address the second issue four years ago and that work is still ongoing.

The potentially productive instructional improvement strategies that we identified constitute an empirically grounded theory of action for instructional improvement at scale. We developed this this theory of action by partnering with several large urban school systems in the US. In the first phase of the project (2007-2011), we partnered with four school systems that served a total of 360,000 students. In the second phase of the project (2011-2015), we continued our work with two of the districts for an additional four years. These collaborations provided a context in which we could test, revise, and elaborate the conjectures about potentially productive improvement strategies that comprised our evolving theory of action.

I first clarify the national context in which we worked by discussing aspects of the US educational policy that influenced our work with the leaders of the four educational systems. I then describe how we collaborated with the leaders and outline the process of testing and revising our conjectures. Against this background, I give an overview of the theory of action for instructional improvement at scale that we developed while working with district leaders.

[1] The work reported in this chapter was supported by the National Science Foundation under grants No. ESI-0554535 and No. DRL-0830029. The opinions expressed do not necessarily reflect the views of the Foundation.
2019. In G. Hine, S. Blackley, & A. Cooke (Eds.). Mathematics Education Research: Impacting Practice *(Proceedings of the 42nd annual conference of the Mathematics Education Research Group of Australasia)* pp. 1-14. Perth: MERGA.

Background: The US Educational System

The US educational system is decentralised and there is a long history of local control over schooling. Each US state is divided into a number of independent school systems or school districts. In rural areas, districts might serve less than 1,000 students whereas some urban districts serve more than 100,000 students. In the US context, urban districts are the largest jurisdictions in which it is feasible to design and implement comprehensive sets of supports for improving the quality of instruction (Supovitz 2006).

The role of the US federal government in education has been quite limited historically when compared with most other industrialized countries. However, in 2001, the US Congress passed a national policy called the “No Child Left Behind” act. Under this legislation, states were given financial incentives to design and implement mathematics standards for student achievement, tests to assess whether students were attaining the standards, and mechanisms for holding schools accountable for increasing scores on those tests. However, the tests developed by most states emphasized procedural skills at the expense of conceptual understanding and problem solving (Shepard 2002). The results of these assessments are consequential in that schools that fail to produce adequate gains in student achievement are sanctioned and, if necessary, reconstituted with a new principal and teaching staff. Not surprisingly, most school and districts responded by “teaching to the test”. As a result, procedurally-oriented assessments came to drive a reform that was intended to focus on rigorous content standards (Resnick & Zurawsky 2005). It was within this national policy context that we recruited four partner districts that were among the minority that aimed to support students’ development of both conceptual understanding and procedural fluency by improving the quality of mathematics instruction.

Clarifying the Problem of Large-Scale Instructional Improvement

In approaching our collaborations with the districts, we assumed that teacher professional development (PD) designed to support teachers' learning would be essential, but we also anticipated that PD would not by itself be sufficient. Findings from a number of prior studies had indicated that the influence of teacher professional development on classroom practice is mediated by the school and district contexts in which teachers work (Blumenfeld et al. 2000; Cobb et al. 2003; Coburn 2003; Spillane 2005). Aspects of these contexts include the instructional materials to which teacher have access, the people to whom they are accountable and what they are held accountable for (e.g., principals' expectations for the structure and focus of lessons, see Cobb & McClain 2006; Elmore 2004), and formal and informal supports for teachers' improvement of their classroom practices (e.g., regularly scheduled teacher collaborative meetings, teachers' informal advice networks). Based on these prior research findings, we anticipated that instructional improvement at scale would involve reorganising the school and district contexts in which teachers work so that those contexts supported teachers' ongoing improvement of their instructional practices. This supposition in turn implicates the practices of mathematics coaches[2], school leaders, and district leaders; the tools they use in their work; and the organisational routines in which they participate (Cobb & Smith 2008).

[2] Coaches are charged with supporting the learning of mathematics teachers in their schools by working both with groups of teachers and with individual teachers in their classrooms.

Project Overview

As we have noted, the four partner districts were all responding to high-stakes accountability pressures by attempting to support and press for ongoing improvements in the quality of classroom instruction. The types of classroom instructional practices on which the districts focused involved achieving rigorous student learning goals by building on students' current reasoning. The development of ambitious instructional practices of this type requires significant learning for most US mathematics teachers and involves reorganising rather than merely elaborating or extending current practices (Franke, Kazemi & Battey 2007).

Our primary research goal in partnering with the districts was to develop a theory of action comprising: 1) a set of instructional strategies intended to support teachers' (and others') learning, and 2) a rationale that justifies why it is reasonable to expect that these strategies will support instructional improvement (Argyris & Schon 1974, 1978). In preparation for our work with the districts, we developed a set of initial conjectures by drawing on the then current literature in mathematics education, teacher education, and educational policy and leadership. These conjectures were broad in scope and spanned curriculum materials and associated resources, teacher PD, teacher collaboration, school instructional leadership, and district leadership (Cobb & Smith 2008). We organised our work with the districts so that it would constitute a context in which to successively test, revise, and elaborate our conjectures, in the process developing an empirically grounded theory of action.

In consultation with district leaders, we selected approximately six middle-grades schools[3] in each district that were representative of district schools in terms of their capacity for instructional improvement. We then recruited 30 middle-grades mathematics teachers in each district, the school leaders and the mathematics coaches who served those schools, as well as district leaders across several central office units, for a total of approximately 50 participants in each district.[4] The data collected in each district each year included: audio-recorded interviews conducted with all participants; online surveys for teachers, coaches, and school leaders; an assessment of teachers' and coaches' mathematical knowledge for teaching; video-recordings of two consecutive lessons in each participating teachers' classroom; video-recordings of professional development; audio-recordings or video-recordings of teacher collaborative meetings; and a teacher network survey.

Each year, we conducted a data collection, analysis, and feedback cycle in each district that involved 1) documenting the district's improvement strategies, 2) collecting and analysing data to assess how these strategies were playing out in schools and classrooms, and 3) reporting the findings to district leaders together with recommendations about how they might revise their strategies to make them more effective.[5] Analysis of the districts' improvement strategies the following year revealed that they acted on 67% of our recommendations (Henrick, Klafehn, & Cobb, 2018). This gave us an opportunity to investigate the conjectures on which our recommendations were based. The revisions we made in the course of successive data collection, analysis, and feedback cycles informed our theory of action, as did both a series of retrospective analyses that we conducted on the data we collected, and developments in the research literature (Cobb et al. 2013).

[3] Middle-grades schools serve students aged 12 to 14 years old.
[4] During the last four years of the project when we worked with only two districts, we doubled the number of schools and participants in each district to have adequate statistical power.
[5] Detailed descriptions of the data collected and the data analysis process can be found in Cobb et al. (2013), and Henrick, Cobb, and Jackson (2015).

Theory of Action for Instructional Improvement in Mathematics at Scale

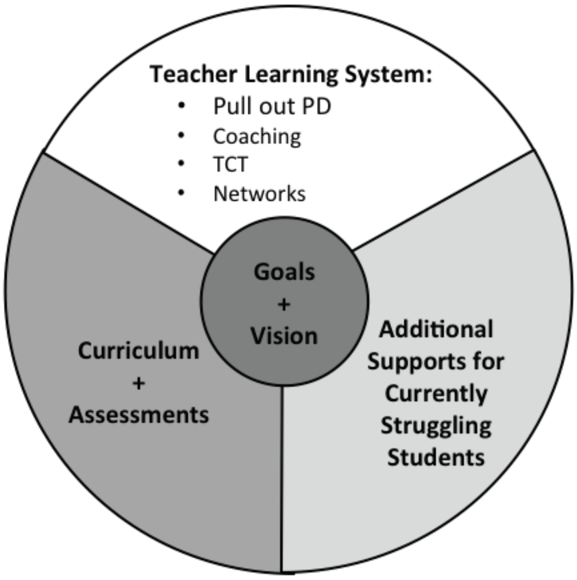

Figure 1: Elements of a coherent instructional system

The theory of action comprises three top-level components: a coherent instructional system[6], school leaders' practices as instructional leaders in mathematics, and district leaders' practices in supporting the development of school-level capacity for instructional improvement. The first of these components, a coherent instructional system, in turn comprises four broad elements: goals for students' mathematics learning and a vision of high-quality mathematics instruction, curriculum and assessments, a teacher learning (sub)system (which includes pull-out PD, school-based teacher collaborative meetings, mathematics coaching, and teachers' informal advice networks), and additional supports for currently struggling students (see Figure 1). Because the responsibilities of U.S. school leaders[7] as instructional leaders differ from those of school leaders in most other countries, I limit my focus in the remainder of this paper to the elements of a coherent instructional system.

Student Learning Goals and Vision of High-Quality Instruction

As shown in Figure 1, goals for students' mathematical learning are at the centre of the instructional system. In the US, mathematics educators have reached a broad consensus on worthwhile learning goals, which include students' development of both conceptual understanding of key mathematical ideas and procedural fluency in a range of mathematical domains (e.g., symbolic expressions, graphs, tables), and learn to communicate their mathematical reasoning effectively by mastering increasingly sophisticated forms of mathematics argumentation (including methods of proof) (National Council of Teachers of Mathematics 2000; National Governors Association for Best Practices & Council of Chief State School Officers 2010). We developed and justified the associated vision of high-quality

[6] The notion of a coherent instructional system based on Newmann et al.'s (2001) seminal work.
[7] We speak of school leaders rather than principals because an assistant principal is sometimes assigned to work with mathematics teachers.

mathematics instruction by drawing on prior research that had identified instructional practices that support students' attainment of these rigorous learning goals (e.g., Kilpatrick, Martin & Schifter 2003; Lester, 2007). As we have indicated, the four districts with which we partnered were attempting to support teachers' development of forms of practice that were compatible with this vision. The vision therefore served to specify goals for teachers' learning and thus oriented their instructional improvement efforts (Jackson, Wilhelm, & Munter, 2018). We delineated the remaining three elements of the instructional system shown in Figure 1 by mapping out from this vision (Elmore, Peterson & McCarthey 1996). In doing so, we drew on our work with the districts, the findings of our retrospective analyses, and the broader literature.

Instructional Materials and Assessments

The second element of a coherent instructional system concerns instructional materials and assessments, particularly formative assessments that can inform the ongoing improvement of instruction. High quality instruction that aims at conceptual understanding and procedural fluency involves engaging students in cognitively demanding tasks that have multiple entry points and can be solved in a range of different ways, thereby enabling the teacher to organise a productive discussion of students' solutions (National Council of Teachers of Mathematics 2000). It is therefore important that the instructional materials that teachers use include such tasks (Stein & Lane 1996). Three of the four districts with which we worked had adopted a textbook series that was aligned with their ambitious goals for students' learning. The fourth district developed its own curriculum and included cognitively demanding tasks.

A retrospective analysis of data collected from the first four years of our work with the districts confirms the importance of teachers using rigorous mathematical tasks as the basis for their instruction. Rosenquist, Wilhelm and Smith (2014) found that the student achievement of teachers who used cognitively demanding tasks was significantly higher than that of teachers who used procedurally oriented tasks, and that the difference was equivalent to between four and six weeks of additional instruction each year. This was the case even though the student assessments focused on procedural competence rather than conceptual understanding and the majority of teachers attempted to "proceduralise" tasks when they introduced them in their classrooms.

Our collaboration with the four districts also highlighted the importance of ensuring that assessments of students' learning align with ambitious student learning goals. Although the districts could not control the nature of state assessments, these assessments did become more aligned with the districts' goals for students' learning during the latter years of our partnership. During these years, the districts also attempted to support teachers' development of formative assessment practices (Black & Wiliam 1988). We found that although teachers did analyse student data during collaborative meetings, they usually did so to identify concepts that they needed to reteach. However, we also found that a significant number of teachers in one district analysed classroom data in order to determine how they could improve their instruction practices (Garner, Kochmanksi, & Henrick, 2018). These teachers first analysed student work in order to assess their learning and then accounted for that learning by relating it their instruction, in the process developing aspects of instruction that they could improve. Significantly, these teachers had participated in pull-out PD, teacher collaborative meetings, and one-on-coaching that were tightly coordinated and constituted a coherent teacher learning subsystem.

Teacher Learning Subsystem

The vision of high-quality instruction located at the centre of the instructional system constitutes the goal for teachers' learning. The third element of the system is a subsystem of supports for teachers' development of ambitious instructional practices. As we have noted, the development of such practices involves significant learning for most US teachers. Our analysis of video-recordings of the participating teachers' instruction indicates that this was the case for our partner districts.

There is growing evidence that, at least in the US context, that pull-out PD alone is frequently insufficient for instructional improvements at scale that require teachers' significant reorganisation of their current practices (Coburn 2003; Elmore 2004). Our partner districts all attempted to support teachers' learning by providing PD, mathematics coaching, and regularly scheduled time for mathematics teachers to collaborate during the school day. However, we found that these supports were usually not coordinated and differed in their intent and focus (Jackson, Horn, Cobb, 2018). For example, one-on-one coaching in teachers' classrooms typically did not build on either the PD in which the teachers had participated or their work during teacher collaborative meetings. In the course of our partnerships with the districts, we came to see the importance of systematically coordinating these supports so that teachers' efforts to improve particular aspects of their practice in one setting either built on or foreshadowed their work in the other settings (Jackson & Cobb 2013).

Pull-out professional development

In the US, districts often organise pull-out PD for teachers from multiple schools by grade level. It is usually designed and led by either the district mathematics specialists[8] or by an external contractor. Prior research indicates that PD has the potential to support instructional improvement provided:

- It is sustained over time and involves the same group of teachers working together (Darling-Hammond et al. 2009).
- It is close to practice (Ball & Cohen 1999), focuses on issues central to instruction, and is organised around the instructional materials that teachers are using in their classrooms.
- It involves both pedagogies of investigation and pedagogies of enactment (Grossman et al. 2009). Pedagogies of investigation entail analysing and critiquing representations of practices, such as video-cases of teaching, in order to develop an image of high-quality practice and to analyse current practice. Pedagogies of enactment involve planning for and trying out intended forms of practice with the support of a facilitator who is an accomplished mathematics teacher (Forman 2003; Lave & Wenger 1991).
- It is led by facilitators who are adept at pursuing their agenda for teachers' learning by building on their ideas differentially, highlighting some of their contributions while letting others drop (Borko et al. 2015; Elliott et al. 2009).

All four districts provided significant amounts of teacher PD that was designed and lead by their mathematics specialists. However, our analyses indicated that this PD was frequently not of high quality. As our goal was to investigate and support the development

[8] District mathematics specialists' responsibilities include supporting the learning of mathematics teachers and mathematics coaches.

of the districts' capacity to improve the quality of mathematics teaching, we did not step in by designing and leading teacher PD. Instead, we conducted a design-based research study in which we investigated how to support PD leaders in learning to design and facilitate high quality PD (Jackson et al., 2015). The details of the study are beyond the scope of this paper, but it is worth noting a few key findings. First, the participating mathematics leaders focused increasingly on central aspects of instruction. Second, they began to develop a vision of teacher PD as supporting teachers' development along a progression of increasingly sophisticated forms of practice. However, we also found that the extent to which district mathematics leaders furthered their agenda for teachers' learning by pressing on their contributions differentially did not improve consistently. A retrospective analysis of the data collected in the course of the study resulted in conjectures about how we could improve our design for supporting PD leaders' learning (Wilson, 2015).

This study illustrates that investigations that contribute to our understanding of what it takes to support instructional improvement on a large scale can be relatively modest in scope provided they target a key aspect of system-level capacity for instructional improvement. There is a need for such studies as the research base that can inform the design and implementation of instructional improvement strategies becomes increasingly thin the further one moves out from the classroom.

Teacher collaboration

At various points during our work with them, all four districts required that principals schedule time during the school day for mathematics teachers to work together. The interviews that we conducted with teachers indicated that they did indeed meet to work on instructional issues. This was not always the case when principals scheduled teacher collaborative meetings either before or after school.

We assessed whether teachers' work in the collaborative meetings was likely to support their development of ambitious instructional practices by first considering the types of activities in which they engaged. Following Coburn and Russell (2008), we distinguished between low-depth and high-depth activities. Examples of low-depth activities include sharing instructional materials and coordinating the pacing of instruction so that all the participating teachers are teaching the same topics. Although these and other low-depth activities might enable teacher to refine their current teaching practices, they are unlikely to support them in fundamentally reorganising those practices. Examples of high-depth activities include working on mathematics tasks together in order to identify the central mathematical ideas and analysing student work in order to identify the range of student thinking on which they could build to achieve a worthwhile mathematical agenda. These activities can potentially provide teachers with opportunities to identify significant conceptual goals for students' learning and to use student work proactively to plan future lessons, not merely retrospectively to assess the extent to which prior lessons were successful (Cobb, Zhao & Dean 2009). These advances would represent significant progress in teachers' development of ambitious practices that align with the districts' agendas for instructional improvement.

In addition to assessing the types of activities in which teachers engaged, we also took account of how those activities were enacted during teacher collaborative meetings. A retrospective analysis of teacher collaborative groups that were productive indicates that the teachers connected content learning goals, students' reasoning, and their instructional practice as they engaged in high-depth activities (Horn, Kane, & Garner, 2018). Enactments of these activities that focus solely on students might account for students' failure to learn as expected in terms of characteristics attributed to students (e.g., perceived limitations in their mathematical capabilities and motivation to learn mathematics). Enactments that connect students and content might account for students' failure to learn by proposing that a particular concept is difficult for seventh graders. In contrast, an enactment that connects content learning goals, students' reasoning, and instructional practice might account for students' failure to learn by proposing that the way the teachers had taught a particular concept made that concept difficult for seventh graders to learning.

We found that only a minority of teacher collaborative groups engaged in high-depth activities, and only a small proportion of these groups enacted high depth activities in ways that were likely to support the development of ambitious instructional practices (Horn, Kane, & Garner, 2018). This indicates the importance of ensuring that teacher collaborative groups include a facilitator who can influence both the types of activities in which the group engages and how those activities are enacted. In our view, at least one member of the group should already have developed relatively accomplished instructional practices, should be constituted as leader in the group, and should press other members on high-leverage issues (e.g., asking, "where's the mathematics?" after group members have worked on a mathematics problem together) (Cobb & Jackson, 2011; Horn, Kane, & Wilson, 2015). Teachers whose instructional expertise is respected by other group members and who receive support in learning to facilitate the group effectively would be good candidates for the role of facilitator. However, in three of the four districts, only a small proportion of the teachers had made significant progress in developing ambitious instructional practices. We recommended to these districts that a mathematics coach attend teacher collaborative meetings whenever possible and serve as the facilitator. One advantage of this arrangement is that coaches might connect work during collaborative meetings to pull-out PD sessions, and might also build on this work when they support group members one-on-one in their classrooms. In doing so, they would play a pivotal role in coordinating these supports so that they constitute a subsystem for teachers' learning (Kane, Cobb, & Gibbons, 2018).

Mathematics coaching

In our view, the general rationale for the role of coaches, who are expected to support both groups of teachers as they plan and analyse their instruction and individual teachers as they interact with their students in the classroom, is reasonably compelling. Research on the development of complex professional practices has documented the importance of co-participating in a practice with a more knowledgeable other (Grossman et al. 2009; Lave & Wenger, 1991). Tharp and Gallimore (1988) argued that co-participation with a more accomplished colleague supports learners in ways that language alone cannot do: "the development of common understanding of purposes and meanings of the activity, [and] the joint engagement in cognitive strategies and problem solving are all aspects of interaction that influence each participant". The work of coaches who can act as more accomplished colleagues appears to be crucial in the US context given that significant learning is required for most teachers if they are to develop ambitious instructional practices. This was particularly the case for the three partner districts in which only a small minority of teachers had made significant progress in developing such practices.

The research base on content-specific instructional coaching was relatively thin when we began working with the four districts. For example, few studies had examined the types of activities in which coaches might engage individual teachers and groups of teachers to support improvements in their classroom practice. We therefore conducted a conceptual analysis of the various types of coaching activities reported in the literature to identify those that are potentially productive in terms of teacher learning opportunities (Gibbons & Cobb 2017; Kane, Cobb, & Gibbons, 2018). Because relevant research on coaching is sparse, we drew on the teacher education literature to develop criteria for determining whether the various coaching activities were likely to give rise to significant teacher learning opportunities. This analysis resulted in the identification of four potentially productive activities that involve coaches working with groups of teachers: 1) engaging in mathematics, 2) examining student work, 3) analysing carefully selected video of classroom instruction, and 4) engaging in lesson study. The remaining three potentially productive activities involve coaches working with individual teachers: 5) modelling instruction, 6) co-teaching, and 7) the coaching cycle in which the coach and teacher co-plan a lesson, the coach observes the teachers' enactment of the lesson, and then the coach and teacher meet to debrief after the lesson. Clearly, the warrant for these findings is weaker than we would like as they are based on a conceptual analysis rather than empirical studies that investigated teachers' learning as they engage in the activities with an accomplished coach. Nonetheless, the findings did enable us to make recommendations about coaching to our partner districts that were based on more that anecdotes and unsubstantiated opinions.

In addition, to identifying potentially productive activities, we also attempted to clarify what coaches need to know and do in order to realise the potential learning opportunities when they enact the activities with teachers. Recent work indicates that coaches should be accomplished mathematics teachers who have relatively deep mathematical knowledge for teaching (Borko et al. 2015; Elliott et al. 2009). We built on these studies by conducting a case study of a coach who was skilled at working with individual teachers in their classrooms (Gibbons & Cobb, 2016). In doing so, we were able to identify five coaching practices that accounted for the focal coach's expertise in working one-one-one with teachers: (a) specifying long-term goals for the teachers' learning; (b) assessing teachers' current instructional practices; (c) locating teachers' current instructional practices on a trajectory of teachers' learning; (d) identifying next steps for teachers' learning, and (e) designing activities to support that learning. In addition, we identified two forms of knowledge that were implicated in the coach's enactment of these practices: (a) a sophisticated vision of ambitious mathematics instruction, and (b) an envisioned trajectory of teachers' development of ambitious instructional practice. As these findings are based on a single case, we consider them to be provisional and open to revision in light of future investigations. In the interim, they provide some guidance for districts that are attempting to support coaches in working more effectively with teachers in their classrooms.

As we have made clear, research on coaching in mathematics and in other disciplines is still at an early stage. It is therefore a prime area for future research, including case studies and design research studies that are feasible as dissertation studies. In our view, it will be important to locate these investigations of coaching within the context of a teacher learning subsystem and a more encompassing coherent instructional system, and to justify particular coaching activities in terms of teacher learning opportunities.

Teacher advice networks

Prior research indicates that teacher advice networks can be a key support for teachers' improvement of their instruction (Frank & Zhao 2005). In contrast to the three types of support for teachers' learning that we have discussed thus far, teacher networks cannot be intentionally designed because they are emergent phenomena that are constituted as individual teachers decide to turn to a colleague for advice about instruction. However, there is evidence that district improvement strategies can influence both to whom teachers turn for instructional advice and the types of advice they seek (Coburn & Russell 2008; Penuel et al. 2009). For example, we found that teachers' participation in productive teacher collaborative groups supported the emergence of advice-seeking interactions that had the potential to support teachers' development of ambitious instructional practices (Wilhelm et al., 2018). It also seems reasonable to conjecture that teachers who use a common set of instructional resources and participate together in PD will be more likely to trust and hold each other accountable, and to turn to each other for instructional advice (Cobb & Jackson, 2011).

Researchers who investigate teacher networks distinguish between who teachers select to turn to for advice and the influence that these interactions have on teachers' knowledge and practice. A retrospective analysis of our network data indicates that teachers whose instructional practices were more sophisticated were more likely to seek instructional advice (Wilhelm et al. 2016). We speculate that this could be because instruction becomes problematic for teachers when they begin to move beyond procedurally-oriented practices. This analysis also found that teachers tended to seek advice from colleagues whose students had higher value-added achievement scores rather than colleagues whose instructional practices were more sophisticated. However, a follow-up case study clarified that who teachers turned to for advice was influenced by the extent to which the principal's and coach's visions of high-quality mathematics instruction were aligned (Rigby, Larson & Chen, in press). In particular, teachers turned to colleagues who had higher value-added scores when the principal pressed them to raise test scores. However, they sought advice from colleagues who had developed more sophisticated practices when the principal and coach both pressed them to improve the quality of their instruction.

This finding about who teachers turned to for advice is significant given the finding of another retrospective analysis which indicates that teachers' interactions with colleagues whose instructional practices were more sophisticated supported the development of the advice seeking teachers' practices (Sun et al. 2014). This analysis also found that the level of sophistication of the practices of the most accomplished teacher in a school was related to overall improvement in quality of instruction in the school.

Taken together, these findings extend previous work by indicating that network interactions can be an important support for teachers' development of ambitious instructional practices. The findings also indicate the crucial contributions that teachers whose instructional practices are relatively sophisticated can make to instructional improvement efforts. In our view, a key aspect of school and district instructional leadership involves identifying such teachers and leveraging their expertise.

Supplemental Supports for Currently Struggling Students

This final element of a coherent instructional system emerged as significant only after we had begun working with our partner districts. As part of their responses to high stakes accountability pressures, school in all four districts invested significant resources to identify and provide supplemental support to students who were likely to perform badly on state assessments. This supplemental support typically included the provision of either individual tutoring or additional, second mathematics classes, both of which were almost invariably procedurally oriented and focused on so-called basic skills.

A retrospective analysis of second mathematics classes across all four districts revealed that the second mathematics classes were rarely effective in improving student achievement on state assessments that focused primarily at low-level goals (Schmidt, 2013). This analysis controlled for both students’ prior achievement and their demographic characteristics, and found there was, on average, a negative or statistically insignificant difference in the achievement of students who were in the additional classes as compared to those who were not. In fact, in only four of thirty schools did students with additional instructional time perform better than expected on state assessments, given their prior achievement. Moreover, this analysis revealed that in schools that implemented supplemental classes, there were lower achievement gains than would have been predicted based on students’ prior achievement for *all* students.

The issue of supplemental supports has broad relevance, as some students are likely to be poorly served in most classrooms including those in which the teacher is striving to develop ambitious instructional practices. In formulating recommendations to the four districts, we argued that the primary goal of supplemental supports should be to enable students who were currently struggling to participate fully in and learn from the instruction in their primary mathematics classes. In an effort to ensure that our recommendations to the districts were empirically grounded, we conducted a literature search to identify studies that addressed this issue in the context of rigorous student learning goals. However, we found few studies that were sufficiently robust to risk using them as a basis for district-level improvement strategies (cf. Confrey, 2011; Nomi & Allensworth, 2009). We therefore drew on the analyses of the districts' supplemental supports that we conducted each year. These analyses indicated that tutoring was unlikely to support students' attainment of ambitious learning goals because it was procedurally focused and because the tutors typically received little if any professional development. However, there were some indications that second mathematics classes were more promising when the same teacher also taught both the primary and supplemental mathematics class and when the teacher's goal in the supplemental class was to prepare students for the primary class (Wilson & Kelley, 2018). We therefore recommended that the districts attempt to meet the needs of currently struggling students by assigning the same teacher to the two classes whenever possible, and support the close coordination of the two classes when this was not feasible. We also suggested that districts provide teachers of second classes with professional development and curricular resources that were specific to meeting the needs of currently struggling students. We took care to signal to district leaders that we regarded our recommendations as provisional as there was little research on which we could draw. In our view, there is an urgent need for studies that address this significant gap in the research literature.

Discussion and Conclusion

In the first part of this paper, we gave an overview of the approach we took for partnering educational systems to identify potentially productive instructional improvement strategies. We then described a key component of our theory of action for instructional improvement – a coherent instructional system – and gave particular attention to the teacher learning subsystem. It is important to acknowledge that the theory of action was developed in the US educational context. Some adaptations will therefore be required to take account of how educational systems are organised in other countries. However, we anticipate that the general approach of framing instructional improvement at scale as a problem of organisational learning will prove relevant across national contexts.

The project on which we reported is large, spanned multiple years, and required significant funding. It is therefore important to emphasise that relatively small studies that investigate the development of key aspects of system-level capacity for instructional improvement can also make significant contributions. The field of mathematics education has made significant progress in documenting trajectories of students' learning in specific mathematical domains, clarifying instructional practices that support students' attainment of worthwhile learning goals, and investigating PD models for supporting teachers' development of those practices. However, this work has had only limited impact on typical mathematics instruction in most countries. This is a strong indication that the standard approach of researchers developing innovative instructional materials and models and then handing them over to practitioners for implementation is ineffective. In my view, there is a pressing need for investigations that challenge this seemingly taken-for-granted division of labour between researchers and practitioners by framing large scale instructional improvement as an explicit focus of investigation. I contend that in the absence of such investigations, mathematics education research is unlikely to realise its potential and contribute to improvements in the quality of mathematics instruction for large numbers of students.

References

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

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

Ball, D. L. & Cohen, D. K. (1999). Developing practice, developing practitioners: Towards a practice-based theory of professional education. In G. Sykes & L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco, CA: Jossey-Bass.

Black, P. & Wiliam, D. (1988). Assessment and classroom learning. *Assessment in Education, 5*, 7-74.

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., Jacobs, J., Koellner, K. & Swackhamer, L. (2015). *Mathematics professional development: Improving teaching using the Problem-Solving Cycle and Leadership Preparation models*. New York, NY: Teachers College Press.

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*, 6-33.

Cobb, P., Jackson, K., Smith, T., Sorum, M. & Henrick, E. (2013). Design research with educational systems: Investigating and supporting improvements in the quality of mathematics teaching and learning at scale. In B. J. Fishman, W. Penuel, A.-R. Allen, & B. H. Cheng (Eds.), *Design-based implementation research: Theories, methods, and exemplars. One hundred and twelfth yearbook of the National Society for the Study of Education* (pp. 320-349). Chicago, IL: National Society for the Study of Education.

Cobb, P. & McClain, K. (2006). The collective mediation of a high stakes accountability program: Communities and networks of practice. *Mind, Culture, and Activity, 13*, 80-100.

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). District development as a means of improving mathematics teaching and learning at scale. In Krainer, K. & Wood, T. (Eds.), *International handbook of mathematics teacher education: Vol. 3. Participants in mathematics teacher education: Individuals, teams, communities and networks* (pp. 231-254). Rotterdam, The Netherlands: Sense.

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*, 203-235.

Confrey, J. (2011): What now? Priorities in implementing the Common Core State Standards for mathematics. Paper presented at the Friday Institute for Educational Innovation, Raleigh, NC.

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N. & Orphanos, S. (2009). Professional learning in the learning profession: A status report on teacher development in the United States and abroad. Dallas, TX: National Staff Development Council.

Elliott, R., Kazemi, E., Lesseig, K., Mumme, J., Carroll, C. & Kelley-Petersen, M. (2009). Conceptualizing the work of leading mathematical tasks in professional development. *Journal of Teacher Education, 60*, 364-379.

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

Elmore, R. F., Peterson, P. L. & McCarthey, S. J. (1996): *Restructuring in the classroom: Teaching, learning, and school organisation*. San Francisco, CA: Jossey Bass

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.

Frank, K. A. & Zhao, Y. (2005). Subgroups as a meso-level entity in the social organisation of schools. In L. Hedges & B. Schneider (Eds.), *Social organisation of schools* (pp. 220-224). New York, NY: Sage.

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

Garner, B., Kochmanksi, N., & Henrick, E. (2018). Educational assessments. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 149-158).Cambridge, MA: Harvard Education Press.

Gibbons, L., & Cobb, P. (2017). Focusing on teacher learning opportunities to identify potentially productive coaching activities. *Journal of Teacher Education, 68*, 411-425.

Gibbons, L., & Cobb, P. (2016). Content-focused coaching: Five key practices. Elementary School Journal, 117, 237-260.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E. & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*, 2055–2100.

Henrick, E., Cobb, P. & Jackson, K. (2015). Educational design research to support large-scale instructional improvement. A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Doing qualitative research: methodology and methods in mathematics education* (pp 497-530). New York, NY: Springer.

Henrick, E., Klafehn, A. B., & Cobb, P. (2018). Assessing the impact of partnership recommendations on district instructional improvement strategies. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 209-220). Cambridge, MA: Harvard Education Press.

Horn, I. S., Kane, B. D., & Ganer, B. (2018). Teacher collaborative time: Helping teachers make sense of ambitious teaching in the context of their schools. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 92-112). Cambridge, MA: Harvard Education Press.

Horn, I. S., Kane, B. D., & Wilson, J. (2015). Making sense of student performance data: Data use logics and mathematics teachers' learning opportunities. *American Educational Research Journal, 52*, 208–242.

Jackson, K. & Cobb, P. (2013). Coordinating professional development across contexts and role groups. In M. Evans (Eds.), *Teacher education and pedagogy: Theory, policy and practice* (pp. 80-99). Cambridge: Cambridge University Press.

Jackson, K., Cobb, P., Wilson, J., Webster, M., Dunlap, C. & Appelgate, M. (2015). Investigating the development of mathematics leaders' capacity to support teachers' learning on a large scale. *ZDM, 47*, 93-104.

Jackson, K., Horn, I. S., & Cobb, P. (2018). Overview of the teacher leaering subsystem. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 65-76). Cambridge, MA: Harvard Education Press.

Jackson, K., Wilheilm, A. G., & Munter, C. (2018). Specifying goals for students' mathematical learning and the development teachers' knowledge, perspectives, and practices. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 43-64). Cambridge, MA: Harvard Education Press

Kane, D. B., Cobb, P., & Gibbons, L. (2018). Understanding content-specific instructional coaching. In P. Cobb, K. Jacksom, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 113-134). Cambridge, MA: Harvard Education Press.

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

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

Lester, F. (Ed.) (2007). *Second handbook of research on mathematics teaching and learnin*g. Greenwich, CT: Information Age Publishing.

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

National Governors Association for Best Practices & Council of Chief State School Officers (2010). Common core state standards for mathematics.
Online: http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf (08.04.2016)

Newmann, F. M., Smith, B., Allensworth, E. & Bryk, A. S. (2001). Instructional program coherence: What it is and why it should guide school improvement policy. *Educational Evaluation and Policy Analysis, 23*, 297-321.

Nomi, T. & Allensworth, E. (2009). "Double-dose" algebra as an alternative strategy to remediation: Effects on students' academic outcome. *Journal of Research on Educational Effectiveness, 2*(2), 111-148

Penuel, W. R., Riel, M., Krause, A. E. & Frank, K. A. (2009). Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record, 111*, 124-163.

Resnick, L. B. & Zurawsky, C. (2005). Standards-based reform and accountability. *American Educator, 29*(1), 8-46

Rigby, J. G., Andrews-Larson, C. J., & Chen, I.-C. (in press). Administrators influence on teachers conversations about the problem of teaching mathematics: A longitudinal case study. *Journal of Educational Change*.

Rosenquist, B. A., Wilhelm, A. G. & Smith, T. M. (2014). The IQA observational rubric and teacher value-added: Looking for nonlinearities to inform practice and policy. *Paper presented at the annual meeting of the American Educational Research Association*, Philadelphia, PA.

Shepard, L. A. (2002). The hazards of high-stakes testing. *Issues in Science and Technology, 19*, 53-58.

Schmidt, R. (2013). *Unpacking tracking: The role of instruction, teacher beliefs and supplemental courses in the relationship between tracking and student achievement*. (Unplublished doctoral dissertation), Vanderbilt University, Nashville, TN.

Spillane, J. P. (2005). *Distributed leadership*. San Francisco, CA: Jossey Bass.

Stein, M. K. & Lane, S. (1996): Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*, 50-80.

Sun, M., Wilhelm, A. G., Larson, C. J. & Frank, K. A. (2014). Exploring colleagues' professional influence on mathematics teachers' learning. *Teachers College Record, 116*, 1-30.

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

Tharp, R. G. & Gallimore, R. (1988). *Rousing minds to life*. New York, NY: Cambridge University Press.

Wilhelm, A. G., Cobb, P., Frank, K. A., & Chen, I.-C. (2018). Teacher advice networks. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 135-148). Cambridge, MA: Harvard Education Press.

Wilson, J. (2015). *Investigating and Improving designs for supporting professional development facilitators' learning*. (Doctoral dissertation), Vanderbilt University, Nashville, TN.

Wilson, J., & Kelley, D. (2018). Supplemental supports for currently struggling students. In P. Cobb, K. Jackson, E. Henrick, & T. M. Smith (Eds.), *Systems for instructional improvement: Creating coherence from the classroom to the district office* (pp. 169-179). Cambridge, MA: Harvard Education Press.