

## Learning to Learn From Data: Benchmarks and Instructional Communities

Suzanne Blanc, Jolley Bruce Christman, Roseann Liu, Cecily Mitchell, and Eva Travers

*Research for Action*

Katrina E. Bulkley

*Montclair State University*

This article examines the use of interim assessments in elementary schools in the School District of Philadelphia. The article reports on the qualitative component of a multimethod study about the use of interim assessments in Philadelphia. The study used an organizational learning framework to explore how schools can best develop the capacity to utilize the potential benefits of interim assessments. The qualitative analysis draws on data from intensive fieldwork in 10 elementary schools and interviews with district staff and others who worked with the schools, as well as further in-depth case study analysis of 5 schools. This article examines how school leaders and grade groups made sense of data provided through interim assessments and how they were able to use these data to rethink instructional practice. We found substantial evidence that interim assessments have the potential to contribute to instructional coherence and instructional improvement if they are embedded in a robust feedback system. Such feedback systems were not the norm in the schools in our study, and their development requires skill, knowledge, and concerted attention on the part of school leaders.

As described in detail elsewhere in this issue (Bulkley, Christman, Goertz, & Lawrence, 2010/this issue) the School District of Philadelphia has responded to the national press for increased student achievement and formal accountability systems by implementing a Managed Instructional System (MIS) with the goal of helping teachers learn more about their students' needs and helping schools reach their mandated performance targets. Interim assessments, aligned with the district's mandated core curriculum, are identified by district leaders as key tools for improving instruction (Bulkley, Christman, Goertz, & Lawrence, 2008). In this article, we draw on the qualitative component of Research for Action's multimethod study to examine how school leaders and formal communities of practice, like grade groups, were making sense of the data provided through the Benchmark assessments—as interim assessments are known in Philadelphia—and how they were able to use these data to rethink instructional practice.

In Philadelphia, classroom instruction in Grades 3 through 8 occurs in 6-week cycles: 5 weeks of instruction, followed by the administration of Benchmark assessments. Then teachers analyze Benchmark data and develop instructional responses to be implemented in the 6th week. The

Benchmark system is designed to provide teachers with formative information to guide what they do in their classroom.

In our multimethod study of Benchmarks in Philadelphia, we found that these assessments are most likely to contribute to improved student learning if there is also concomitant attention to developing strong school leaders who promote data-driven decision making within a school culture focused on strengthening instruction, professional learning, and collective responsibility for student success. The quantitative component of the study showed greater achievement growth at schools where there was more evidence of implementation of the *managed instruction system*—that is, greater use of the Core Curriculum, more professional development on use of student data, greater access to technology, and more collegial work on how to respond instructionally to Benchmark data (Christman et al., 2009). However, the most robust predictors of learning growth were instructional leadership, collective responsibility, and use of the Core Curriculum. The qualitative component of the study indicated that practitioners' use of Benchmark data was not as connected with changes in instructional practice as the proponents of interim assessments might hope. However, the qualitative analysis also showed that use of Benchmark data was more likely to be connected to instructional improvement when school leaders focused on developing robust instructional communities that supported teachers in interpreting Benchmark data in the light of the learning goals established by Philadelphia's Core Curriculum.

Organizational learning theorists argue that instructional communities, such as grade groups, provide "an ideal organizational structure for school staff to learn from data and use data to improve student learning" (Mason, 2003, p. 1). But few studies of schools have looked closely enough at collective interpretation of data to see how teachers and school leaders actually engage in making shared sense of data—what they talk about and how they talk about it. This led us to ask how grade groups made sense of data from Benchmark assessments and generated actionable knowledge for planning instructional improvements.

Understanding sense-making and follow-up actions is especially important because interim assessments may be used for multiple purposes. They may *predict* student performance on an end-of-year summative, accountability assessment; they may provide *evaluative* information about the impact of a curriculum or a program; or they may offer *instructional* information that help diagnose students' strengths and weaknesses (Perie, Marion, Gong, & Wurtzel, 2007). Although a primary intention of Philadelphia's leaders was that Benchmarks would provide teachers with information to diagnose and address students' strengths and weaknesses, we found that the uses of Benchmark data did not always align with the District's intended purposes. A focus of the study is the use of student performance data within feedback systems that help practitioners transform data into actionable knowledge. In our effort to understand how Benchmark data contribute to organizational learning, we applied the concept of a four-step "feedback system" to analyze the structures and processes educators use to engage with data collectively and systematically during the course of a school year. The four steps in the feedback system are (a) accessing and organizing data, (b) making sense of data to identify problems and interventions, (c) trying the interventions, and (d) assessing and modifying the interventions.

Figure 1 is a visual representation of this feedback system, a model that builds on the previous work of Richard Halverson and his colleagues (Halverson, Prichett, & Watson, 2007). In examining how this type of system might work to transform Benchmark data into actionable knowledge, we are particularly interested in "microtasks" (e.g., displaying the data, formulating

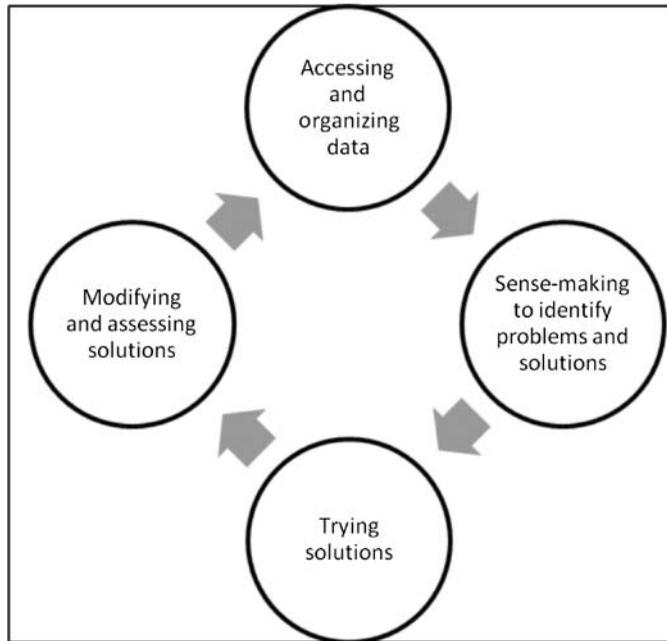


FIGURE 1 Feedback system for engaging with data.

substantive and provocative questions about the data) as well as “macrofunctions” (e.g., encouraging data-driven decision making; Spillane, Halverson, & Diamond 2001).

In this article, we explore the following questions:

- What do school leaders and teachers working in instructional communities talk about as they make sense of Benchmark data? What plans do they make as a result of their interpretation of the data?
- What micropractices of school leaders and instructional communities have the potential to cohere the steps of feedback systems and to enrich the use of Benchmark data?
- What implications does an analysis of the micropractices of school leaders and instructional communities have for professional development that will enable school leaders to make the most of interim assessments?

## OVERVIEW OF THE ARTICLE

In the sections that follow, we present our research methods and conceptual framework. Then we describe and analyze the sense-making process that we observed school leaders and teachers engage in as they examined Benchmark assessment data during grade group meetings. This analysis is followed by a case study of a school with robust feedback systems for using Benchmark data. We end the article with a discussion of some of the ways that school leaders and instructional

communities can mitigate the negative influences that the high-stakes accountability environment has on the use of Benchmark data.

## RESEARCH METHODS

This article is part of a study that was conducted from September 2004 through June 2007. During the first year, the research was exploratory in nature and focused on learning about the district's MIS as it unfolded, identifying schools that exemplified effective use of data, and working with the district to develop and pilot a district-wide teacher survey that included items related to data use. The entire study drew on three kinds of data:

- a district-wide teacher survey administered in the spring of 2006 and 2007
- student-level demographic data and achievement data from standardized tests
- qualitative data obtained from intensive fieldwork in 10 elementary schools and interviews with district staff and others who worked with the schools, as well as further in-depth case study analysis of five schools in 2006–2007

The school sample included 10 elementary schools that were among the 86 schools identified as “low performing” and eligible for intervention under a state takeover of the Philadelphia school district (Christman, Gold, & Herold, 2006). The 86 low-performing schools represented 39% of the district's 220 elementary and middle schools. Like the other 76 low-performing schools, each of the 10 schools in our sample was assigned to an intervention model beginning in the 2002–2003 school year. Seven of the schools were under management by outside providers, 2 schools were part of the district's homegrown intervention under the Office of Restructured Schools, and 1 school was a “sweet sixteen” school—a low-performing school that was showing improvement and therefore received additional resources for two years but was not under a special management arrangement. An original intention of the study was to use the different management models as points of comparison. However, this research purpose fell away when all of the provider organizations, except Edison Schools, Inc., adopted the district's MIS.

In identifying schools to be part of the qualitative study, we sought out schools from each intervention model that would provide insight about how schools learn to engage with data as part of a process of school change. We developed a purposive sample of schools that were identified by district staff, provider staff, and other school observers as being well on the road to making effective use of data. Criteria used for identification included the following data-driven decision making was a stated priority of school leaders; professional development on how to access, organize, and use Benchmark data was ongoing; and grade group meetings to discuss Benchmark data occurred regularly.

All of our schools served a considerably higher percentage of students living in poverty than the district average and served student populations that were predominantly either African American or Latino (Christman et al., 2009). In 2005–2006, a team of at least two researchers made two one-day site visits to each of the 10 schools. During each visit, we conducted semistructured interviews with the principal and two or three teacher leaders. Interviews lasted 60 to 90 min. Site visits were scheduled on days when we also could observe a leadership team meeting, grade group meeting(s), or other data related event(s).

In 2006–07, we narrowed our sample to five schools for more intensive fieldwork. To select the 5 schools, we developed criteria related to 4 categories: the principal’s role in data use, the strength of the professional community, the school’s adequate yearly progress (AYP) status, and the purposes that school staff brought to their interpretation of Benchmark data. The research team placed schools along continua for each category and selected schools that represented the range of variation. Two researchers spent about 4 days in each school. During these visits, we followed up with principals, observed several events at which staff discussed data, talked extensively with teacher leaders, and interviewed at least two classroom teachers in each school. By June 2007, our qualitative data set included more than 150 interviews with school staff and faculty; 54 observations of leadership team meetings, grade group meetings, and school-wide professional development sessions; and a collection of school documents.

To analyze the interviews, we coded the data using Atlas.ti and identified themes and practices within and across schools using content analysis. We used information from written documents and field observations to triangulate our findings. Other analytical strategies included case study write-ups of data use in each of the 10 schools, reduction of data into word charts (e.g., a chart describing the types of data that were attended to by school staff, the settings and actors involved, and the resulting instructional decisions), and the development of extended vignettes of feedback systems in schools.

## Organizational Learning and Organizational Linkages

The importance of an organizational feedback system for data use derives from organizational theorists’ argument that actionable and meaningful “knowledge is most likely generated when people work together in tightly knit groups” (Brown & Duguid, 1998, p. 28). According to organizational learning theory, learning organizations are composed of interconnected communities of practice. Within the communities, workers are engaged in a *joint enterprise*, they are bound together in *mutual relationships* characterized by common norms for participation and performance, and they employ a *shared repertoire of habits of mind* that involves them in thinking through problems of practice (Brown & Duguid, 2000; Lave & Wenger, 1991; Wenger, 1998).

Much of the research on instructional communities has focused on identifying and measuring the *qualities*—such as openness to new ideas, trust, collaboration, shared responsibility for student learning—that lead to improved outcomes for students (i.e., Kruse, Louis, & Bryk, 1995). Literature on organizational learning directs us to a *process* perspective on instructional communities by emphasizing the importance of *routine practices* within and across these communities, for example, pointing to the importance of *tools* (e.g., meeting agendas, protocols for examining student work) and *organizational structures* (e.g., regular times set aside for meetings) used by teachers and leaders as they work together (Cobb, McClain, Lamberg, & Dean, 2003; Lave & Wenger, 1991; Supovitz & Christman, 2003; Wenger, 1998). Research that has examined the organizational structures, roles, and objects that promote shared practices within a school also shows the value of collaboration among educators. According to Kruse and his colleagues,

Individual skills and knowledge provide the foundation of the school’s capacity, but a school’s ability to manage complex cycles of innovation depends on the ingrained habits of learning from colleagues both within and across work groups. (Kruse et al., 1995, p. 34)

In this article we look at linkages such as tools, people, and structures that have the potential for helping practitioners in diverse roles (e.g., classroom teachers, teacher-leaders, administrators) develop shared understandings of the goals and strategies that can cohere and guide them across various settings and activities (e.g., daily classroom instruction, periodic professional development activities, annual school-wide planning efforts). From an organizational learning perspective, if a school is characterized by strong linkages, then teachers and administrators will have a shared framework for interpreting a range of activities as part of a common endeavor and will be more likely to undertake actions that support innovation and improvement. Borrowing the language from the community of practice literature, linkages can be thought of as “boundary objects,” “boundary brokers,” and “boundary encounters” that promote shared learning and span the work of people in a variety of different roles, settings, or groups (Brown & Duguid, 1998; Wenger, McDermott, & Snyder, 2002). Turning to our examination of how schools can learn to learn with data, we focus on the tools, people, and structures that have the potential to “link” varied activities and individuals into a coherent feedback system for using Benchmark data to improve student learning.

### *Tools*

The tools associated with Philadelphia’s MIS are designed to create an integrated system of instruction and assessment that coordinates the work of individual teachers, schools, and central office staff. The Core Curriculum and its associated standards are at the center of this system as described in detail by Bulkley and others. The accompanying planning and scheduling time line structures the content and concepts taught across classrooms and across grade levels. Together, the Core Curriculum and the planning and schedule timeline create instructional guidelines that are aligned with state content standards and that prepare students for the state tests.

Benchmark data reports are intended to scaffold the implementation of the Core Curriculum by providing interim information to teachers, students, parents, and administrators on student performance during each 6-week Benchmark cycle with information about student strengths and weaknesses on material that is taught during each 6-week Benchmark cycle. Classroom teachers typically look at reports that indicate students’ proficiency levels and their responses to each Benchmark item. The district’s online data system (SchoolNet) also provides the opportunity for teachers and school leaders to generate a variety of other reports that track the progress of individual students over time, generating information about the success of grades, schools, or other groupings, or information about the success of individuals or groups in particular content strands or standards. Thus, the MIS is a tool that creates a linkage between the district’s curricular goals, its performance goals, and the actual performance of any individual or group within the system.

### *School Leaders*

A district’s curricular tools can provide guidance to a school, but principals play a critical role in the quality of a school’s academic program (Bryk, Sebring, Kerbow, Rollow, & Easton, 1998). Principals are also important in creating an environment in which adult learning is central to school improvement (DuFour, 2002; Elmore, 2000). In looking at instructional communities

as communities of practice, we also take a view of leadership that goes beyond the principal. We follow the reasoning of Spillane et al. (2001), among others, who have argued for a conception of leadership that is *distributed* across the roles of principal, teachers, and parents/community members.

School leaders are likely to play a major role in creating linkages in feedback systems for Benchmark data because interpretation of information and its context is a key function of organizational leaders (Daft & Weick, 1984). The transformation of Benchmark data into actionable knowledge requires interpretation, including interpretation of the goals for using these data. Any boundary object (such as Philadelphia's Core Curriculum or Benchmark Analysis Report) that moves across different communities of practice will be interpreted and perceived in different ways by the different communities (Cobb et al., 2003; Wenger, 1998). As previously noted, Philadelphia's MIS is designed to provide high-quality instruction, but it is also designed to be aligned to the state's testing system. Furthermore, as we noted, the possible uses of interim assessments include uses such as predicting student success on year-end tests, evaluation of the implementation of programs and curricula, as well as formative feedback for instructional change (Perie et al., 2007). Because of the possibility for many purposes and goals for interim assessments, it cannot be assumed that staff within a school will have a shared understanding of the purposes of Benchmarks. Nor can it be assumed that school-level staff will have the same understanding of these goals as district-level leaders. In this context, school-level leaders must play the role of "boundary brokers" who help staff make sense of the district's stated expectation that Benchmarks will be used for instructional improvement as they move toward shared norms for interpreting and acting on Benchmark data. In addition, school leaders are in a position to help their schools and staff learn from data by providing follow-through for instructional decisions made on the basis of Benchmark data, providing support to ensure the quality of these interventions, and using additional data to assess and refine the interventions implemented on the basis of Benchmark results.

### *Structured Encounters*

In examining how schools learn to learn with data, we were particularly interested in exploring how grade group meetings become communities of practice that develop shared knowledge and practice related to Benchmark data. Given the importance of shared tasks and processes within our theoretical framework, our analysis of grade group meetings focused on their routine processes, with special attention to the distinction between "macrofunctions" and "microtasks." To identify the microprocesses, we carefully documented the timing, setting, talk, and structures of grade group meetings. We also looked at the tools that were used for organizing and making sense of data within grade group meetings and at the roles of leaders and participants in these meetings in accessing and organizing data, interpreting data, and developing strategies and plans for follow-up. Detailed field notes from these meetings allowed us to examine and assess the quality of interpretation processes and the factors that influenced that quality.

### **Making Sense of Benchmarks: What Happens in Grade Groups?**

In this section of the article, we draw on our observations of grade group meetings to make visible the purposes that school leaders and teachers brought to their interpretation of Benchmark data and the instructional actions that they took based on their interpretations.

Our observations of grade groups suggest that practitioners engaged in three major types of sense-making as they sat together to discuss and interpret Benchmark data: *strategic*, *affective*, and *reflective*. Not surprisingly, the pressures of the accountability environment strongly influenced their sense-making because of the pressure to meet AYP. However, our observations also showed that the actions of school leaders could mediate these policy forces to create instances of substantive professional learning for school staff. Disappointingly, such instances occurred less frequently and less powerfully than one would hope. Next we discuss the three kinds of sense-making. Although we discuss each type of sense-making separately, they are often intertwined in grade group conversations.

*Strategic sense-making* focuses on short-term tactics that help a school reach its AYP targets. Strategic sense-making includes conversations about “bubble students” who have the highest likelihood of moving to the next level of performance (from below basic to basic or from basic to proficient) thereby increasing the probability that the school will meet its AYP goal. These conversations relate to the predictive purpose of interim assessments in Perie et al.’s (2007) framework, described earlier. Strategic conversations also focus on improving test-taking conditions and test-taking strategies. Finally, in strategic conversations, practitioners use Benchmarks for evaluative purposes as they work to identify strengths and weaknesses that cut across grades and classrooms so that they can allocate resources (staff, materials, and time) in ways that increase the odds that the school will meet its AYP goal (e.g., assigning “strong” teachers to the accountability grades).

In our observations, strategic sense-making dominated the talk about Benchmark data. School leaders borrowed terminology from the state accountability system, applied it to the Benchmark results, and opened meetings with the questions, “How many students are Proficient or Advanced? How many are close to Proficient or Advanced? What are the questions that gave the students the most problems?” Even though the Benchmark data are meant to provide *diagnostic* information about what students have learned in the previous 5 weeks, conversations about results often assumed that they were *predictive* of performance on the Pennsylvania State System of Assessment (PSSA)—evidence of how the state’s accountability measure pervaded practitioners’ thinking about what they could learn from the Benchmark data. Practitioners from all of the schools in our qualitative sample reported that the identification of bubble students—students on the cusp of scoring proficient or moving from below basic to basic—was a common practice in their analysis of Benchmark data. One example from our teacher interview data shows the kinds of practices associated with strategic sense-making:

We have a chart, and what we did is we tracked. . . . The teachers put stars next to those kids that they’re going to target. And we made sure that those kids had interventions, from Saturday school to extended day, to Read 180, which is another intervention. And then we followed their Benchmark data. But those were the kids that the teachers were really going to focus on, making sure that those kids become proficient, or move that 10 percent out of the lower level so that we can make Safe Harbor next year.

*Affective sense-making* includes instances in which leaders and classroom teachers address their professional agency, their beliefs about their students, their moral purpose, and their collective responsibility for students’ learning. During affective talk, school leaders and teachers offer one another encouragement. They express a “can do” attitude, often relating this sense of

professional agency back to the pressures that they feel from the accountability environment. In affective talk, practitioners also affirm their belief that their students “can do it” —whether that means moving to proficiency on the PSSA test or mastering a skill in the Core Curriculum. They discuss how to motivate their students to put forth their best effort on standardized exams. This description of one grade group meeting helps to demonstrate affective sense-making in practice:

At this school, the 5th-grade teachers said that their students were having a lot of problems with fractions on the Benchmark tests, particularly reducing improper fractions. But one teacher explained that she had connected fractions to a lesson that she had done earlier with her class and that, “A lot of light bulbs went off [when students saw how to draw on what they already knew].” Building on this, the principal said that she loved the image of students drawing on what they already knew and suggested that everyone make posters of light bulbs for their classroom to motivate students during the Benchmarks and other tests.

In this example, the principal diverted the conversation away from content and pedagogy to address how to motivate students. She encouraged teachers to help their students believe they “can do it”—an example of affective sense-making in which school-based practitioners focus on how to encourage their students. Notably, no one in the meeting addressed conceptual issues related to mathematical content. Students were challenged by items related to fractions, but the conversation did not explore the intended purpose of these questions. Similar to Spillane and Zeuli’s (1999) findings in their study of mathematics reform, our research indicates that discussions about Benchmark data most often did not focus on building teachers’ “pedagogical content knowledge” (Shulman, 1987). Pedagogical content knowledge couples knowledge about content to knowledge about pedagogy. Teachers with strong pedagogical content knowledge understand what teaching approaches fit the content being taught; their deep understanding of content makes it possible for them to explain disciplinary concepts to students and to craft learning tasks that build students’ conceptual understanding; their broad repertoire of instructional strategies provide them with options to help students with different learning needs. The alignment of Benchmark assessments with the Core Curriculum offers the opportunity for teachers to look at results with an eye toward strengthening their pedagogical content knowledge. Our observations of grade group meetings and our interviews with school leaders indicate that this was rarely a focus of practitioners’ analysis.

*Reflective sense-making* occurs when teachers and leaders question and evaluate the routine instructional practices that they employ in their classrooms and their school. They connect what they are learning about what their students know and do not know to key concepts in the Core Curriculum and identify resources that will help them strengthen instruction of those concepts and tailor it to the needs of individual students and groups of students. These purposes relate to both the instructional and evaluative purposes described by Perie et al. (2007). Researchers have pointed out the importance of reflective discourse as “a springboard for focused conversations about academic content that the faculty believes is important for students to know” (Mintz, Fairman, & Buffet, 2006, p. 94). These connections help teachers and leaders to focus on what they need to learn in order to help their students succeed. At one school—which is discussed in depth later in this article—grade group meetings became a routine opportunity for teachers to connect Benchmark data systematically back to the performance standards of the Core Curriculum. Such discourse about the curriculum serves to shift teachers’ attention away from students’ failures and toward analyzing and strategizing about their own practices.

During reflective talk, teachers and leaders also compare Benchmark data to data from other assessments and from their own observations of students. Another excerpt of field notes of a grade group meeting shows how leaders can model reflective sense-making in instructional communities:

The math teacher leader suggested that middle grade students need more independence during regular classes in order to improve their performance on tests. “One of the reasons that people say the kids know the material, but don’t test well, is that the conditions are so different. During instructional periods, you need to let the kids do more on their own, so it’s more like a testing situation where they have to interpret the instructions on their own.” He suggested that the teachers should tell students the objective for the lesson, then have them work in small groups to figure out what is being asked of them in the directions for the math activity. Teachers should circulate during this time, noting where students are on the right track and where they are not. They should ask questions that will help students improve their interpretations. He concluded, “Our students need to learn to be more independent. After they’ve finished the task, then you can review and reflect with the small groups about how it went.”

In this example, the math leader makes the broad point that students need to learn to work more independently. She then offers specific ideas for doing this. This example indicates how different kinds of sense-making intertwine; although these suggestions are meant to address problems students encounter in the testing situation, they are also good instructional practice.

Offers of support from school leaders were prominent in our observations of grade group meetings, as were teaching tips. Principals and teacher leaders offered to conduct demonstration lessons and to consult about classroom management of small groups. They also suggested steps that teachers might take themselves. We read these offers of support and recommendations as ways for school leaders to demonstrate their investment in teachers’ struggles and to encourage teachers in the context of the larger accountability policy context that stigmatizes schools, educators, and students for low student achievement rather than supporting and rewarding them.

Our interviews of staff suggest that follow-up by principals and teacher leaders in classrooms was much less likely to occur than one might hope, a gap that weakens the kinds of feedback systems necessary for organizational learning. When leaders do not visit classrooms to see whether teachers are trying the strategies discussed in grade group meetings and whether they use the strategies well, an important evaluative function of Benchmark assessments is lost. Leaders do not have good information to judge the efficacy of the solutions that teachers are trying.

In summary, reflective conversations can help practitioners plan the kinds of professional development that will strengthen teachers’ understanding and use of the Core Curriculum. They catalyze school leaders and teachers to think about what other kinds of data they need to consider as they make sense of the Benchmark results. We also hypothesize that these kinds of conversations will lead to an increase in school and classroom instructional capacity.

Just as there were many types of talk about Benchmarks, there were many types of plans made on the basis of Benchmark data. These plans include strategies explicitly geared to making AYP, strategies for addressing gaps in student knowledge, strategies for dealing with affective aspects of learning and testing, and strategies for deepening the knowledge and skills of teachers. A consequence of reviewing Benchmark data, analysis of our observational data indicate that practitioners planned actions that included the following:

1. *Identifying students who were likely to move from “Basic” to “Proficient” or from “Below Basic” to “Basic” and targeting them for special interventions to increase the*

*likelihood that the school will make AYP.* Across the schools, these interventions varied considerably—extended day programs, Saturday school, work with volunteer tutors, special attention from the math or reading specialist, computer-assisted programs. It is likely that their quality varied as well, but formal or informal assessment of the interventions was rare. As one principal told us, “You know, we’ve never really looked to see if those tutors are doing a good job.”

2. *Identifying skills and concepts to be retaught in the 6th week of the instructional cycle or in subsequent units.* From our data, we surmise that reteaching was one of the actions most frequently taken as a result of reviewing the Benchmark results. District leaders and principals reported that there were many instances of teachers simply returning to the content material and using the same instructional strategies. But some teachers reported that it was important to try different instructional strategies for reteaching an area of weakness. As one explained,

I can see how my whole class is doing. And they [members of my grade group] can say, “This one question, only four of your twenty kids got it right.” So, I know that if only four kids got it right, that’s something I need to go back and re-teach, or get a fresh idea about how to give them that information.

3. *Identifying students who shared similar weaknesses (or, in some cases, strengths) for regrouping to provide differentiated instruction.* Our data indicate that regrouping was also one of the actions most frequently taken as a consequence of reviewing the Benchmark results. Often referred to as “flexible groupings,” teachers and school leaders explained that they grouped students around shared weaknesses identified through examination of the Benchmark data. One teacher described how “the groups constantly changed” so that she could “target specific kids and their specific needs and group kids according to where they were lacking.” When she felt it was appropriate, she would also assign different homework to different students based on their needs. In other schools, teachers described how they had begun creating groups that cut across classrooms based on shared student weaknesses.
4. *Rethinking classroom routines that emphasized greater student independence, motivation, and responsibility for their own learning.* This action was discussed infrequently. For example, a fifth-grade teacher described how she regrouped students, putting stronger students with weaker students as a way to encourage and facilitate peer teaching.

I put the item analysis on the overhead projector [for the whole class to see]. It’s because of that relationship I have with my students. It’s that community. So [I want my students thinking about] why is our class average 60% when I scored 100%, I didn’t get any wrong. We need to help our classmates that had difficulty, that may have received 40%. That’s where I go into my grouping. How can I pool my strong students [to work with students who are struggling]?

5. *Identifying content and pedagogical needs of teachers to inform opportunities for continued professional learning and other supports that addressed those needs.* Formal professional development sessions and less formal on-the-spot coaching were also planned based on results from the Benchmarks, especially when those data corroborated data from

the PSSA. One teacher described a particularly strong approach to supporting teachers' learning:

We actually had a professional development about [guided reading] where [the principal] did a lesson to show us, and then we went to two other teachers' rooms and saw them do a lesson. And then pretty much that whole week that followed, [the principal] came around to see how we were using it, if we needed any help, what other support we needed to get this going and into play.

Each of these planned actions makes sense. Each emerged from paying attention to data. However, the quality of the actions varied considerably. Spillane, Reiser, and Reimer (2002) argued that educators' interpretations of policy mandates are critical to their implementation of these mandates. In the previous examples, we note the influence of the accountability environment on educators' interpretation of the mandate for data-driven decision making. Clearly, this policy context and the fact that these schools had been identified as "low performing" influenced practitioners' perceptions of why examining data is important. They needed to address the primary problem that they felt compelled to solve: how to make AYP. They brought the imperative to "do something"—some might say "do anything"—to their discussion and interpretation of Benchmark data.

If practitioners focus only on superficial problems—described as "the low-hanging fruit" by principals in our study—their intervention strategies are likely to be mundane. As Sarason (1982) and others have shown, schools have often been adept at "adaptive learning" that maintains rather than interrupts the status quo when new information, ideas, and innovations are introduced. Beer and Eisenstat (1996) lay out the significance of organized talk to organizational learning:

Lacking the capacity for open discussion, [practitioners] cannot arrive at a shared diagnosis. Lacking a shared diagnosis, they cannot craft a common vision of the future state or a coherent intervention strategy that successfully negotiates the difficult problems organizational change poses. In short, the low level of competence in most organizations in fashioning an inquiring dialogue inhibits identifying root causes and developing fundamental systemic solutions. (pp. 599–600)

Our data indicate that the quality of the sense-making in instructional communities determines the quality of the actions that practitioners take based on the data.

### How Micropractices of School Leaders Capitalize on Student Performance Data: The Case of the Mahoney Elementary School

In this section, we focus on the Mahoney Elementary School to look in detail at an example in which leaders, tools, and encounters shape strong processes for collective learning from Benchmark data. Under the leadership of Mrs. Sylvia Bannon, teachers at Mahoney helped their students move from overwhelming failure to a point far beyond the requirements of No Child Left Behind. Before Mrs. Bannon's tenure the majority of students routinely scored "below basic" on standardized tests. During the period of our research, students scored more than 80% proficient and advanced in mathematics and 60 to 70% advanced or proficient in reading ([www.portal.state.pa.use/portal/server.pt/community/school/assessments/7442](http://www.portal.state.pa.use/portal/server.pt/community/school/assessments/7442)).

During the years of Mrs. Bannon's tenure at the Mahoney School, collective examination of Benchmark data at Mahoney was integrated into a strong *feedback system* in which teachers were able to look at data, reflect on their own instructional strengths and weakness, and experiment with new practices designed to better meet the needs of their students and the expectations of the school's leadership. When Mrs. Bannon left Mahoney, proficiency rates dropped back down to approximately 50%, suggesting that the leadership transition was difficult for the school even with some stability among other members of the leadership team. Nevertheless, we believe the example has much to offer to others who are interested in learning how school leaders can help build strong instructional communities for engaging with Benchmark data.

At Mahoney, instructional change was supported as school leaders followed up in classrooms to help teachers with new instructional practices and to modify these practices where appropriate. As in other schools, the interpretation of Benchmark scores included a significant amount of strategic talk about passing the PSSA. However, at Mahoney, engagement with the Benchmark data that included all three kinds of talk was also an important component of *individual and organizational learning*. The case study of Mahoney illustrates how a talented and committed school leader fostered a school-wide approach to Benchmark data as part of her overall commitment to using student data as a lever for improvement in a school serving one of Philadelphia's most impoverished African American communities. Mrs. Bannon, the principal of Mahoney and a member of a family of African American educators, established the educational expectation and moral imperative that every child in the school could and would learn to be a skilled reader, writer, problem solver, and critical thinker. She conveyed this commitment in her interviews with us, in her comments at grade group meetings, and in her approach to school-wide improvement.

In our final interview with her in the spring of 2007, Mrs. Bannon reflected on her commitment to addressing the challenges of African American students as well as her belief in the value of external accountability. Talking about the school's professional development activities, she told us that teachers were reading Alfred Tatum's 2005 book *Teaching Reading to Black Adolescent Males* as one text in a series of voluntary breakfast meetings.

Later in the interview Mrs. Bannon referred to her commitment to African American students as an explanation of her support for No Child Left Behind.

There shouldn't be any children left behind. We must focus on that 20% or that 18% or that 32% because they don't deserve to be left behind. We all need to be accountable for what we do. The pressure from the tests creates an atmosphere of learning and ensuring that kids are making it and are going to achieve well on a test they know they've got to take.

She ended her discussion of No Child Left Behind by reading a passage from Alfred Tatum's book about the need to help young African American males develop self-determination by helping them to connect with texts that are meaningful to them. For Mrs. Bannon, teacher accountability and empowered young African Americans went hand-in-hand.

Analysis of interviews transcripts and field notes collected over a 2-year period indicate a remarkable level of consistency in the ways that staff and leaders at Mahoney described the school's shared norms and routines. The principal's expectations created discomfort for some staff members. However, the principal's commitment to children was acknowledged and respected. Although participation in collective activities was required rather than voluntary, staff involvement with Benchmark data became part of a joint enterprise for developing a shared repertoire of habits of mind (Brown & Duguid, 2000; Lave & Wenger, 1991; Wenger, 1998).

In the rest of this section, we explore how the encounters, the tools, and the leaders at Mahoney positioned Benchmark data within strong feedback systems to strengthen teaching and learning throughout the school at the same time that Benchmark data were utilized as a strategic tool for addressing accountability requirements.

### *Structured Encounters*

Grade group meetings were a key site at Mahoney. These were held weekly and included the principal, the math teacher-leader, the literacy teacher-leader, and the two or three classroom teachers for each grade. These were described by the principal and teacher leaders as the most important site in the school for teacher learning. In fact, during the second year of our research, Ms. Bannon told us that grade groups were renamed “Professional Learning Communities” to highlight their contribution to teachers’ professional development.

Grade group meetings at Mahoney were highly structured and consistently focused on instructional issues. The weekly grade group meetings began with a member of the leadership team handing out a typed agenda with a guiding question at the top, ended with the principal summarizing next steps, and were followed up with typed notes distributed to all participants. The guiding questions were not necessarily profound, but they did help to ensure focus and coherence to the weekly meetings. For example, in the meeting discussed next, the guiding question was a restatement of common questions about Benchmark assessments: Where did student achievement improve? Where did it decline?

According to teachers and school leaders, grade group meetings always centered on analysis of data or reflection on instruction. As one teacher told us, “Everything begins by talking about data.” In addition to providing opportunities for engaging with Benchmark data, grade group meetings were opportunities for teachers to move cards indicating students’ reading levels and report on student progress in reading, to share and assess student writing, and to discuss other instructional issues. For example, toward the end of our second year of research, Ms. Bannon used the fourth-grade meeting as an opportunity to ask the fourth-grade teachers to reflect on what they had done well, to reflect on what they could do differently next year, and to prepare a presentation for the rest of the faculty about their own goals for improvement. As in other schools, Mahoney’s grade group discussions of Benchmarks encompassed the three types of interconnected sense-making: strategic, affective, and reflective. We discuss each of these in turn.

The analysis and interpretation of Benchmarks at Mahoney not only was focused on instruction but also highlighted their connection to high-stakes tests. Like other principals, Mrs. Bannon treated the Benchmark assessments as if they were predictors of success on the PSSA, although they were not designed for this purpose, and the scoring is in fact different. As in observations at other schools, Mrs. Bannon asked, “How many students are Proficient or Advanced? How many are close to Proficient or Advanced?”

As at other schools, talk about Benchmarks and the PSSA also led into talk about the school’s moral purpose and leaders’ belief in the capabilities of their staff and students. Whenever students scored below proficient, Mrs. Bannon reminded the staff that this was not acceptable and that teachers were always responsible for changing their own behavior to reach all of their students. As she said at one grade group meeting to a teacher whose students were not improving their

reading skills, “You don’t have the option. [The children] need high-quality instruction. They need planned, sustained support. These children do not have other options.”

The expectation that all students would be proficient was accompanied by a consistent focus in grade groups on the Core Curriculum, the standards, and what teachers could do to improve their own teaching. This focus was consistent across all grade levels. For example, in the area of literacy, we saw a strong focus on ensuring that students acquired early literacy skills in the primary grades and moved into content-rich reading and thinking in the upper elementary and middle school grades. As one teacher said,

The school has been focused on using the data to help the kids and push the instruction. Every kind of thing that we do, every assessment we give, we look at it; we see what we need to change, and how we can differentiate our instruction so that it’s helping them do more.

### *Tools*

At Mahoney, multiple tools were utilized to help coordinate the work of teachers and leaders within the school. Some tools were directly related to the Benchmarks. Other tools, notably the Core Curriculum, were also part of the MIS. Mahoney had still another set of tools that were not directly related to the analysis of Benchmark data but were nevertheless critical for developing and implementing a school-wide approach to instructional improvement. These included formats for common lesson plans and guidelines for teacher leaders as they visited classrooms. As in other schools, Benchmark Item analysis reports (the tool that provides information about individual students’ answers to each question as well as summaries for the entire class) were important tools in grade groups. However, at Mahoney, the Core Curriculum was another key tool in grade group meetings. Teachers were expected to bring the curriculum framework to grade group meetings so they could examine it and refer to it at the same time that they discussed the standards in which their students’ showed weaknesses. In addition, in preparing for grade group meetings at Mahoney, teachers used the District’s Benchmark Data Analysis Protocol to assess the weaknesses in their students’ Benchmark results, identified strategies for improving the areas of weakness, and prepared to share these strategies with their colleagues.

The literacy teacher explained how conversations at grade group meetings helped teachers understand the connections between the Benchmark reports, the Data Analysis Protocol, and the Core Curriculum:

When we get results back, we always highlight weaknesses and strengths. When we focus on the weaknesses, that’s where the Core Curriculum comes into play. [We tell the teachers,] “If you’re struggling, turn to the [Core Curriculum] picture window.”

Teachers also articulated how grade group meetings helped them understand the standards that are the basis of the Philadelphia Core Curriculum as well as the PSSA test. According to one teacher, “If you say ‘standard 1.1,’ people might not know what it is. [At these meetings, we talk about] what does it look like in the classroom?”

The Mahoney principal and teacher leaders identified several other tools that were important in developing shared instructional routines and strategies in the school. One key tool was a required format for the 90-minute mathematics and literacy blocks in the school. These had been developed

collaboratively by the teacher leaders and classroom teachers several years before our study and were still in use. For example, the format for the math lesson identified how many minutes were to be spent on each element of the lessons specified in the *Everyday Mathematics* curriculum, such as “math boxes” for quick review of skills, activities to introduce new concepts, and “math games” designed to provide opportunities for students to work with skills and concepts. Teacher leaders also used a tool that they had adapted from an observation format used by district coaches who had worked with them previously. This tool helped to guide their classroom visits and to identify aspects of classroom instruction such as use of centers, differentiation of instruction, and fostering of higher order thinking. A third tool was the lesson plan that was turned in by every teacher and reviewed by the principal.

### *School Leaders*

As previously discussed, Mahoney’s principal, Mrs. Bannon, played a strong leadership role in establishing norms and practices in the school. Distributed leadership was one of the important practices that Mrs. Bannon advocated and implemented. Mahoney’s math teacher leader and reading teacher leader were both fully released from regular classroom instruction. They worked with teachers to identify and implement short-term interventions based on Benchmark data, they helped teachers understand what Benchmark information was most important to pay attention to and what Benchmark information could be ignored, and they collaborated with the principal on developing long-term strategies for meeting the school’s goals.

Ms. Bannon explained why she had prioritized putting limited resources into full-time teacher leaders when she became the Mahoney principal a few years before our study started.

It was a hard decision since it meant larger class sizes. But I wanted to begin with a strong leadership team. And it’s a choice between having a great teacher reach 25 students or having a great teacher reach other teachers.

The varied contributions of the teacher leaders at Mahoney were apparent in both interviews and observations. For example, in a grade group meeting that we observed, the math teacher leader

- pointed out that using calculators would improve student scores on a significant number of questions
- offered to do a workshop for teachers about how to use their classroom sets of calculators as part of the upcoming professional development day
- explained that “matrix multiplication” showed up on the Benchmarks, but was a technique that is specific to a particular curriculum and wouldn’t be on the statewide tests
- offered to come into the fourth-grade classrooms and do demonstration lessons on expanded notation for small groups of students.

The teacher leaders cohered the stages of the feedback system first by helping teachers make sense of Benchmark data, then by providing support to teachers as they tried out new teaching practices, and finally by helping teachers assess and revise their strategies. Because the teacher leaders were in and out of the classrooms, they could assess whether teachers were successfully

changing their instructional practices. As members of the leadership team, and as participants in grade group meetings, the teacher leaders could continue to work with the principal and the teachers to modify interventions and identify new interventions as needed.

Many teachers interviewed, especially in the lower grades, articulated the value of the teacher leaders' ongoing support. One teacher said, "Knowing that my literacy leader is there, and if I say to her, 'You know, I'm not really sure how I'm going to do this lesson,' she's always there and very helpful." Staff also reported that at the subsequent grade meetings, there were chances to discuss how well strategies identified at the previous grade groups had been working.

### Organizational Learning and Instructional Coherence

It is important to note that the processes described in the grade group meeting just presented were typical of meetings at Mahoney. Meetings typically had a formal agenda, included some examination of assessment data, discussions of a content area needing improvement, and reflections about what could be done differently. Because these meetings were usually attended by Mrs. Bannon and the teacher leaders, the knowledge generated in any one part of the school (e.g., an effective instructional strategy or a new interpretation of a key Core Curriculum concept) could be shared with teachers in other groups and classrooms. During Mrs. Bannon's tenure at Mahoney, the Benchmarks became a powerful vehicle for reinforcing the use of the curriculum, for focusing teachers' attention on the standards, and for organizing conversations about ways in which teachers could improve their teaching. Because these school-based discussions were framed by the standards and strategies embedded in Philadelphia's Core Curriculum, they helped nurture the "*instructional coherence*" identified by the Consortium for Chicago School Research as showing a positive impact on student learning (Newmann, Smith, Allensworth, & Bryk, 2001).

Research for Action's quantitative study indicated that instructional leadership, collective responsibility, and use of the core curriculum were strong predictors of learning growth. The quantitative analysis, like the qualitative analysis, indicates that there is something important about relationships among school leaders and faculty involving accountability to each other and use of the core curriculum that support student learning. It is notable that the measures of school leadership and school community are stronger predictors of student learning growth than technology, or professional development, or satisfaction with Benchmarks. Although these aspects of a MIS are helpful, and maybe even essential, they are not in themselves sufficient to bring about increases in achievement without a community of school leaders and faculty who are willing and able to be both teachers and learners.

The Mahoney case study provides some insight into how instructional leadership, collective responsibility, and the use of the Core Curriculum might create a context in which Benchmarks contribute to growth in student learning. The case study suggests that the 6-week Benchmark cycle (with its connection to the curriculum and standards, its regularly timed feedback to teachers, and its tools for analysis and reporting) is potentially a powerful tool for supporting instructional coherence and for making visible the skills needed to implement a coherent instructional program. However, the qualitative analysis also suggests that if the conditions for using Benchmarks to promote instructional coherence are not present, Benchmark data will be used in superficial ways, if at all. The principal at Mahoney saw the potential in this system and took advantage of it by structuring Benchmark discussions as one of many opportunities for teachers to think

and learn more about what is important in the Core Curriculum, how to use the curriculum, and how to get their students to higher levels of proficiency. Ms. Bannon also devoted resources, including her time and that of the two teacher leaders, to prioritizing classroom coaching and professional development that would ensure that all teachers would have the skills and knowledge to consistently implement the district's curricular framework.

In our earlier discussion of how educators make sense of Benchmarks, we suggested that short-term strategic uses of Benchmark data often take precedence over other purposes. At Mahoney, the principal and teacher leaders established a set of structures and practices that ensured that Benchmark data were used as part of a process for ensuring high-quality instruction. They created an organization in which teachers, as well as students, were constantly in learning mode—and their success was reflected in the significant achievement gains of their students.

## CONCLUSION

The use of interim assessments by large urban school districts to improve instruction and student achievement is on the rise. The findings from this research on how practitioners use these assessments in Philadelphia's K–8 schools offer formative lessons about the importance of formal instructional communities as settings for school-based practitioners to learn from interim assessment data so that they can make substantive improvements in their teaching. Next we discuss the implications of this research for district and school leaders in Philadelphia and elsewhere. The research also has important implications for the higher education community that educates and certifies district and school leaders across the country.

The most important message from this research is that the success of even a well-designed system of interim assessments is dependent on the knowledge and skills of the school leaders and teachers who are responsible for bringing the system to life in schools. Stringent accountability measures, strong curricular guidance, and periodic assessments are not substitutes for skilled and knowledgeable practitioners working together in instructional communities to use data to improve instruction. Investments in human capital cannot be bypassed. Data can make problems more visible, but only people can solve them. Mandated accountability measures are insufficient for instigating the kind of collegial relationships that result in shared responsibility for school improvement and improved student learning.

In Philadelphia, immediate needs for improved testing outcomes pressed practitioners to use Benchmark data to predict student performance on the PSSA—a purpose that Benchmarks were not designed to serve. These needs often worked against instructional communities capitalizing on their discussions of Benchmark results as opportunities to learn more about how to help all students master the concepts and skills of Philadelphia's Core Curriculum.

However, our research also indicates that the use of Benchmark data is not always a narrow exercise in preparing to “teach to the test.” In the language of organizational learning, school leaders and instructional communities enacted organizational practices that contributed to *individual* teacher learning and professional growth while fortifying a *collective* understanding of the challenges, goals, and path ahead for the school.

Looking more broadly, organizational learning offers a robust framework for Understanding what school leaders need to know and be able to do in order to make formal instructional communities a robust site for professional learning from interim assessment results and other kinds

of data about student achievement. As we have argued, interim assessment data will contribute to changes in teaching and learning only if it is situated within a feedback system in which practitioners access and organize data, interpret data to identify problems and solutions, try out solutions in their classrooms, and modify their solutions based on new assessments. As learning leaders, principals and teacher leaders need deep content knowledge and strong facilitation skills to lead the kinds of deliberative conversations that make discussions of interim assessment results opportunities for teacher learning. School leaders must ensure that the interpretation processes in instructional communities are rife with opportunities for teachers to question their understanding of the pedagogy embedded in curricular frameworks and instructional materials. They need to learn how to frame conversations about assessment data so that teachers understand the connections to larger school improvement priorities and to the school's curriculum. They need to know how to pose questions in ways that invite teachers to talk openly about curriculum concepts, how their students learn best, what instructional practices have worked and those that haven't, what additional curricular resources they need, what they need to learn about content, and where they might seek evidence-based instructional strategies that would address the learning weaknesses of their students. They also need to know how to steer teachers away from inappropriate use of data to predict performance on standardized tests. School leaders need opportunities to develop this knowledge base and practice these skills. Understanding the value and purposes of the different types of sense-making identified in our research—*affective, strategic, and reflective*—and how to use them offer a framework for such training.

As learning leaders, principals and teacher leaders need to know how to allocate resources and establish school organizational structures and routines that support the work of instructional communities and assure that the use of interim assessment data is embedded in the feedback systems necessary for organizational learning. School schedules need to accommodate regular meetings of grade groups. Principals and teacher leaders need to be at these meetings and, with teachers, establish meeting routines that include agendas, discussion protocols with guiding questions, and documentation of proceedings. Follow-up to the meetings is crucial. School leaders need to visit classrooms to see if and how teachers are using instructional strategies and to offer resources and coaching so that teachers can deepen their understanding of curriculum content and pedagogy. Assessing the impact of interventions is crucial. Important steps include helping teachers to design classroom-based assessments for use during the 6th week of instruction, encouraging peer observations, and examining the quality of common interventions such as tutoring and after-school remediation programs. School leaders must recognize their role in the creation and diffusion of knowledge across the school. There are also lessons from the research and an organizational learning framework that have implications for districts might consider as they design and implement their own periodic assessments. As districts and schools develop organizational structures, processes, and tools to support the use of interim assessment data, they need to ask themselves a number of questions: Are these structures, processes and tools supporting the review of data as a *collective activity of instructional communities*? Are they supporting the review of data as an activity that helps teachers deepen their pedagogical content knowledge? Are they supporting the multiple steps of feedback systems, especially the assessment and modification of interventions?

This analysis of the micropractices of school leaders and instructional communities has revealed many steps district and school leaders can take to realize the value of interim assessments—and data from other assessments—as tools for learning. The analysis in this article

provides substantial evidence that interim assessments can contribute to instructional coherence and instructional improvement when they are embedded in a robust feedback system. However, the analysis also confirms a common concern: In the absence of leadership and instructional communities that are able to engage deeply and reflectively with interim assessment data, interim assessments can all too easily—and incorrectly—be positioned as activities that provide simple information that will help students and schools succeed within an environment of high-stakes accountability.

## REFERENCES

- Beer, M., & Eisenstat, R. A. (1996). Developing an organization capable of implementing strategy and learning. *Human Relations*, 49, 597–619.
- Brown, J. S., & Duguid, P. (1998). Organizing knowledge. *California Management Review*, 40(3), 28–44.
- Brown, J. S., & Duguid, P. (2000). Organizational learning and communities of practice: Toward a unified vision of working, learning, and innovation. In E. L. Lesser, M. Fontaine, & J. Slusher (Eds.), *Knowledge and communities* (pp. 99–121). Oxford, UK: Butterworth-Heinemann.
- Bryk, A. S., Sebring, P. B., Kerbow, D., Rollow, S. G., & Easton, J. Q. (1998). *Charting Chicago school reform: Democratic localism as a lever for change*. Boulder, CO: Westview.
- Bulkley, K. E., Christman, J. B., Goertz, M., & Lawrence, N. (2008, April). *Building with Benchmarks: The role of the district in Philadelphia's Benchmark assessment system*. Paper presented at the annual meeting of the American Education Research Association, New York.
- Bulkley, K. E., Christman, J. B., Goertz, M. E., & Lawrence, N. R. (2010/this issue). Building with Benchmarks: The role of the district in Philadelphia's Benchmark assessment system. *Peabody Journal of Education*, 85(2), 186–204.
- Christman, J. B., Bulkley, K., Neild, R., Blanc, S., Liu, R., Mitchell, C., et al. (2009). *Making the most of interim assessments: Lessons from Philadelphia*. Philadelphia: Research for Action.
- Christman, J. B., Gold, E., & Herold, B. B. (2006). *Privatization "Philly style": What can be learned from Philadelphia's Diverse Provider Model of school management?* Philadelphia: Research for Action.
- Cobb, P., McClain, K., Lamberg, T., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher*, 32, 13–24.
- Daft, R. L., & Weick, K. E. (1984). Towards a model of organizations as interpretation systems. *Academy of Management Review*, 9(2), 284–295.
- DuFour, R. (2002). The learning-centered principal. *Educational Leadership*, 59(8), 12–15.
- Elmore, R. F. (2000). *Building a new structure for school leadership*. Washington, DC: The Albert Shanker Institute.
- Halverson, R. R., Pritchett, R. B., & Watson, J. G. (2007). *Formative feedback systems and the new instructional leadership* (WCER Working Paper No. 2007–3). Retrieved from <http://www.wcer.wisc.edu/publications/workingPapers/index.php>
- Kruse, S. D., Louis, K. S., & Bryk, A. S. (1995). An emerging framework for analyzing school-based professional community. In K. S. Louis & S. D. Kruse (Eds.), *Professionalism and community: Perspectives on reforming urban schools* (pp. 23–42). Thousand Oaks, CA: Corwin.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Mason, S. A. (2003, April). *Learning from data: The role of professional learning communities*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Mintz, E., Fairman, S. E., & Buffett, T. (2006). Digging into data. In K. P. Boudett, E. A. City, & R. J. Murnane (Eds.), *Data wise: A step-by-step guide to using assessment results to improve teaching and learning* (pp. 81–96). Cambridge, MA: Harvard Education Press.
- Newmann, F. M., Smith, B., Allensworth, E., & Bryk, A. S. (2001, January). *Improving Chicago's schools: School instructional program coherence: Benefits and challenges*. Chicago: Consortium on Chicago School Research. Retrieved from [http://ccsr.uchicago.edu/content/publications.php?pub\\_id=39](http://ccsr.uchicago.edu/content/publications.php?pub_id=39)

- Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007, November). *The role of interim assessments in a comprehensive assessment system: A Policy brief*. Washington, DC: The Aspen Institute.
- Sarason, S. B. (1982). *The culture of the school and the problem of change*. Boston: Allyn & Bacon.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
- Spillane, J. P., Halverson, R. R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher*, 30(3), 23–28.
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72, 387–431.
- Spillane, J. P., & Zeuli, J. S. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis*, 21(1), 1–27.
- Supovitz, J. A., & Christman, J. B. (2003). *Developing communities of instructional practice: Lessons from Cincinnati and Philadelphia*. Madison, WI: Consortium for Policy Research in Education.
- Tatum, A. (2005). *Teaching reading to black adolescent males: Closing the achievement gap*. Portland, ME: Stenhouse.
- Wenger, E. (1998). *Communities of practice: Learning, meaning and identity*. New York: Cambridge University Press.
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice*. Boston: Harvard Business School Press.

Copyright of PJE. Peabody Journal of Education is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.