American Educational Research Journal
April 2015, Vol. 52, No. 2, pp. 208–242
DOI: 10.3102/0002831215573773

Making Sense of Student Performance Data: Data Use Logics and Mathematics Teachers' Learning Opportunities

Ilana Seidel Horn
Britnie Delinger Kane
Jonee Wilson
Vanderbilt University

In the accountability era, educators are pressed to use evidence-based practice. In this comparative case study, we examine the learning opportunities afforded by teachers' data use conversations. Using situated discourse analysis, we compare two middle school mathematics teacher workgroups interpreting data from the same district assessment. Despite similarities in their contexts, the workgroups invoked different data use logics that shaped teachers' learning opportunities. The first workgroup's instructional management logic linked increasing student achievement to individualization. The second workgroup's instructional improvement logic focused on students' thinking and linked it to instructional changes but was limited by broader instructional management logics. Evidence-based practice cannot be understood apart from the data use logics in teachers' communities, which are shaped by policy constraints.

KEYWORDS: accountability policy, case study, data use, discourse analysis, teacher learning

In the context of No Child Left Behind (NCLB), accountability pressures are shifting the ground of teachers' work, especially in frequently tested

ILANA SEIDEL HORN is an associate professor of mathematics education at Vanderbilt University's Peabody College of Education and Human Development, PMB 230, GPC, 230 Appleton Place, Nashville TN 37203-5721; e-mail: *ilana.s.horn@vanderbilt.edu*. Her research focuses on teachers' situated learning and equitable mathematics teaching practices.

BRITNIE DELINGER KANE is a doctoral candidate at Vanderbilt University's Peabody College of Education and Human Development. Her research uses discourse analysis to study pre-service and in-service teacher learning about ambitious instruction.

JONEE WILSON is a doctoral candidate at Vanderbilt University's Peabody College of Education and Human Development. Her research interests are professional development and equitable instructional practices in secondary mathematics.

subjects like mathematics. Performance data have become a focus for teachers and administrators, particularly in schools not meeting achievement benchmarks—adequate yearly progress (AYP) in the policy's terms (Diamond & Cooper, 2007). Given this increased attention to performance data, it is important to understand the policy's influence on instruction. As others have pointed out, NCLB is the current instantiation of our nation's penchant for standards-based educational reform (Mehta, 2013). In his historical analysis of the role of standardization in U.S. education, Mehta summarizes the diverging views of standards over time: one view emphasizes competition and accountability through external pressure, while the other seeks to professionalize by developing shared knowledge and skill among practitioners.

NCLB falls into the former standardization camp, as an external accountability policy that sets desirable outcomes through proficiency cuts and pass rate targets on standardized tests. In this way, NCLB's logic resonates with notions of productivity and accountability akin to those of business and industry. Frequent high-stakes testing—and the related rankings and sanctions of students, teachers, and schools—provides the mechanism for external accountability and standardized outcomes, yet preferred methods of instruction remain underspecified. Because of this ambiguity, the press to improve education via NCLB has led, in some quarters, to schools "improving" not so much by supporting teachers' professional learning but by reorganizing instructional work—with or without fundamental change in instructional practice—in the service of improved standardized test scores (Diamond & Cooper, 2007). As shorthand, we refer to this version of standardization as indicating an *instructional management logic* (Jackson, Cobb, & Rigby, 2014).

In contrast, Mehta's latter form of standardization emphasizes the role of professionalism and sees the development of pedagogical skill as central to success. This version of standardization emphasizes improving educational outcomes by supporting teachers' professional growth. Accountability systems should provide technical assistance to teachers, allowing best practices to rise to the top, while valorizing the professional knowledge of teaching. Although Mehta argues that NCLB drives toward the instructional management logic, in some quarters, schools and districts are trying to leverage it as a mandate for professionalization, investing in teachers to increase their knowledge, skill, and ultimately, effectiveness. We refer to this interpretation of accountability as *instructional improvement logic* (Jackson et al., 2014). Obviously, these logics are not mutually exclusive. As Jackson and colleagues describe, they require tight coordination so that the short-term management goals do not entirely overshadow long-term improvement goals.

NCLB has reawakened these endemic tensions in discourses about teaching, learning, and school change, as educators negotiate instructional management and instructional improvement logics at all levels of the

enterprise. These tensions often surface in interpreting and using student performance data. Because they work in a frequently tested (and frequently troublesome) subject, mathematics teachers are regularly asked to use and respond to such data. For these reasons, mathematics teachers' data use practices provide an instructive site to examine the policy's influence, as this group has often been in the hot seat of accountability pressures. In particular, we take mathematics teachers' data use conversations as useful occasions to understand the relationship between accountability policy, data use logics, and ultimately, teachers' learning about instruction.

Data Proliferation and Data Use in Schools

Questions about data use in schools grow ever more urgent as performance data proliferate beyond annual testing. Because assessments mandated by NCLB do not provide feedback until the end of the academic year, districts often turn to interim assessments designed to give teachers guidance about their current students' performance, leaving schools "data rich" and "information poor" (Stringfield, Wayman, & Yakimowski, 2005). These interim assessments are typically formatted like high-stakes tests (Supovitz, 2012). Although they are marketed as formative assessments, they function primarily for predictive purposes, which Wiliam and Thompson (2007) have described as "early warning summative tools" (p. 7). The increased frequency of testing and its related anxieties make data conversations commonplace in many schools.

Data use consists of a wide-ranging set of practices that continue to multiply in schools, not all of which conform to an instructional improvement logic (Ikemoto & Marsh, 2007). First, we know that data alone do not "speak" but rather require interpretive work to be meaningful (Honig & Coburn, 2008). Raw data must be transformed into "usable knowledge" for educators to apply them to subsequent action (Mandinach, Honey, Light, & Brunner, 2008). In fact, educators need substantial assistance to effectively use data (Wayman & Stringfield, 2006; Wohlstetter, Datnow, & Park, 2008). Finally, the mere presence of data (even usable data that educators have been trained to use) does not guarantee its contribution to instructional improvement. Issues such as timeliness, perceptions of validity, and alignment of data use practices to other educational goals have been identified as critical for data to meaningfully inform instruction (Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006). Complicating matters further, data use is not just one practice: different forms of data may require different practices (Schildkamp & Kuiper, 2010).

Spillane's (2012) call to understand data use in schools' everyday practice is most relevant to the present analysis. He argues that the particulars of data use are consequential to how policies like NCLB ultimately influence changes in practice, with particular attention to the organizational routines

that shape their use and interpretation (Feldman & Pentland, 2003; Sherer & Spillane, 2011). We focus our lens a grain-size down from organizational routines to teachers' conversations. This leverages an analysis of teachers' learning by examining the relationship between *particular data use practices* and *teachers' opportunities for learning*. Incorporating insights from studies of discourse and learning, we address pressing questions about what kinds of data, in what forms, in what types of activities, might illuminate meaningful issues in teaching and contribute to instructional improvement. This analysis also responds to Little's (2012) call to examine data use from a microprocess view. By analyzing meaning(s) constituted in teachers' interactions, we uncover how data use contributes to teachers' learning. As analysts, we attempt to put aside our pedagogical values and remain agnostic about the endpoint of teachers' learning. In other words, teachers may be learning practices that we might find counter to good instruction; as researchers of everyday learning, we view these as examples of learning nonetheless. Through this lens, we aim to contribute an understanding of how schools and policies create teacher-learning environments, for better or for worse, with the hope that they can be more deliberately organized to improve instruction.

Teachers' Conversational Learning With Data

In this section, we present our theoretical perspective on teachers' conversational learning and explore how data use might shape these interactions. After briefly introducing the learning theory that informs our work, we focus on the idea of *learning opportunities*. Then, we specify how such learning opportunities arise in teachers' conversations. Finally, we consider the role that data can play in these conversations, setting the stage for our analysis.

Studying Learning Opportunities in Teachers' Conversations

From a sociocultural perspective, learning involves changing participation in a community of practice (Lave & Wenger, 1991), a framework that highlights the social resources for and consequences of people's learning. In this light, we investigate teachers' learning in data conversations by analyzing *learning opportunities* therein. Specifically, we identify how activities and environments stand to support new forms of practice and understanding (Greeno & Gresalfi, 2009). This emphasis on "opportunities" foregrounds the group process over the particular impact on individuals. To operationalize this construct, we look for how conversations (a) provide teachers with conceptual resources and (b) mobilize teachers for future work (Hall & Horn, 2012; Horn & Kane, in press). In other words, through an analysis of conversational content and processes, we uncover concepts communicated and their implications for teachers' future instruction.

This perspective on conversational learning opportunities builds on Horn and colleagues' prior work. This program of research includes comparisons of mathematics teachers' talk and learning in different school settings (Horn, 2005, 2007), to teachers in other subject areas (Horn & Little, 2010; Nolen, Horn, & Ward, 2011), to the workplace talk of other professionals (Hall & Horn, 2012), and among mathematics teachers at different levels of instructional accomplishment (Horn & Kane, in press). In sum, this research points to the way everyday workplace conversations both reveal and influence locally acceptable notions of teaching, mathematics, and student learning, affording teachers particular learning opportunities while limiting others, serving as a *conceptual infrastructure* for the work of teaching (Horn, 2005).

Relevant to the present analysis, these studies point to features of teachers' talk that shape and are shaped by local workgroup cultures and contribute to learning opportunities. These features vary across settings, partly due to available expertise but also because of differences in workgroup norms, patterns of interaction, resources, and values, which, together, distinguish work cultures (Van Maanen & Barley, 1984). Because these microcultural differences can be difficult to pin down, we parse teachers' discourse to identify their manifestations. Figure 1 represents the conversational features in workgroup conversations that reflect and reinstantiate these cultures. Although meaning develops across these features, we disaggregate them for analytic purposes. We describe each conversational feature in the next section, linking each one to its contribution to meaning making and, in turn, to teachers' learning opportunities.

Making Meaning in Workgroup Conversations: Data Use in Interaction

In this section, we explain how teachers' talk renders aspects of their work for joint consideration, with an eye on how it shapes learning opportunities. Before explicating the conversational features in Figure 1, we draw attention to another key analytical object in our study of data use: the data report. The distinct "voice" of data reports requires an extension of prior work on teachers' conversational learning. To account for the reports' agency in data conversations, we borrow a concept from actor network theory (Latour, 2009) and view data reports as durable aggregates of social forces that "speak" in teachers' conversations. Endowing inanimate objects with agency may seem like a peculiar analytic move, but the data reports exert power in workgroup conversations as teachers contend with their narratives about teaching and learning. The social forces of the accountability system gather in the reports, making them difficult for educators to ignore. While reports, as objects, do not have volition like human actors, they are nonetheless important interlocutors. As we describe below the components

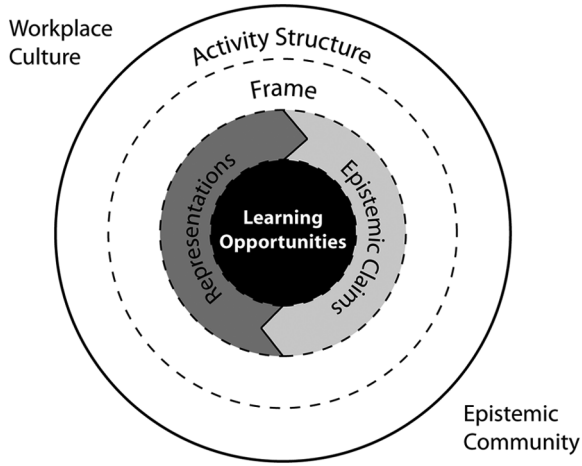

Figure 1. **A schematic representation of the conversational features that supported sensemaking in teacher workgroups. These resources work in concert to shape meanings and professional learning opportunities as they support images of future desirable practice.**

of Figure 1, we consider the data reports' interface with each conversational feature.

Epistemic Stances

The first conversational feature that shapes meaning in teachers' workgroup talk is what we call *epistemic stances* (Hall & Horn, 2012; Horn & Kane, in press). These refer to workgroups' perspectives on what can be known, how to know it, and why it is of value. Analytically, teachers' conversations reveal epistemic stances on different timescales of interaction. At the smallest time scale, bald declarations in single turns of talk (e.g., "what matters here is motivating the kids") manifest these stances, what we call *epistemic claims*. At a slightly broader time scale, a question implies an epistemic stance (Horn & Little, 2010). For instance, the question, "Do you think the kids fell apart because they were tired?" conveys a stance that *sometimes fatigue can perturb student behavior*. Over longer time periods, epistemic stances surface through activities, as interactional emphases and attention reveal commitments to what can be known and how to know it, while providing different interpretive resources on the work.

Data reports, with their sociotechnical associations, also convey epistemic stances. Their scientific veneer—computer printouts of numbers, statistics, and distributions—often results in their conversational positioning as

objective measures of student learning, rather than one of many possible accounts of it. Data reports effectively make epistemic claims about who is and who is not learning, without a humble sense that the test is only one means for interpretation. As data reports find their way into teachers' conversations, so too do test-related category systems—like the *commended, passing, bubble, growth* scheme derived from NCLB (Booher-Jennings, 2005)—become meaningful lenses for seeing instructional effectiveness. The introduction of data reports thus shifts the epistemological grounds for interpreting teaching and learning from a collection of warranted inferences to definitive interpretations buttressed by testing technologies. While the epistemic stances carried by data reports do not necessarily supersede the workgroup's, the reports' accounts are made powerful through accountability policy and thus need, at the very least, to be negotiated.

Representations of Practice

The second conversational feature that contributes to teachers' sensemaking is *representations of practice* (Little, 2003). Teachers typically discuss their work asynchronously from active instruction, so workgroup conversations require participants to reconstruct critical aspects of teaching to ground their discussions. That is, workgroups cannot talk productively unless they are discussing more or less the same thing, and representations provide a basis for intersubjectivity. Horn (2005, 2010) identified two forms of teacher talk that served to interactionally reconstruct critical classroom events for joint consideration. *Replays* are stories that provide blow-by-blow accounts of classroom interactions or ongoing events, while *rehearsals* allow teachers to act out interactions in a more general or anticipatory fashion.

Because testing events are the basis for data reports, these matter in teachers' sensemaking, extending replays and rehearsals beyond critical events of classroom teaching to encompass critical events of testing days: the student who came late, where to place a timer in the room, and the like. These representations support teachers' sensemaking in data use.

Data reports themselves also constitute important representations of practice. Unlike more familiar artifacts like curriculum or student work, data reports represent stories that have material consequences for school personnel and are thus more likely to shift a workgroup's focus. As representational tools, data reports' particular form influences their contributions to teacher sensemaking (Mandinach et al., 2008). They necessarily highlight certain facets of teaching while obscuring others. For instance, reports disaggregated by topic provide different conversational traction than lists of summative student scores. Similarly, reports focusing on student subpopulations, with no mention of content, provide different leverage still. For this reason, we relate representational details of data reports to teacher learning.

Activity Structures

The third conversational feature that contributes to teachers' sensemaking is *activity structures*. This refers to the patterned ways tasks get carried out in group interaction. To qualify as an activity structure, tasks may be formalized or improvised as people work, or some combination of the two. Over time, repeated activities may become routine—predictably patterned and recurrent (Horn & Little, 2010). While such *conversational routines* share features of Feldman and Pentland's (2003) organizational routines, with an interplay of predictability and improvisation, our analysis focuses at the level of interaction rather than the organization.

Activity structures and representations often work together in teachers' data use. For example, in the activity of looking at student work, teachers might all look at the same work samples or they may look only at their class sets. These differences in activity require different representations, which are consequential for teachers' learning opportunities. In the former case, they can delve into a subset of solutions to investigate students' thinking together. In the latter case, teachers' knowledge of students can be brought to their interpretations of the work.

Problem Framing

The final conversational feature that contributes to teachers' sensemaking is *problem frames*. Closely related to activity structures, problem frames describe how issues are defined through interactions (Goffman, 1974). Frames organize teachers' collective attention, shifting the meaning of activities (Horn, 2007), differentially positioning teachers as agents in problems they face (Bannister, 2015), thus changing the nature of learning opportunities. To understand the connection between frames and epistemic stances, consider the activity *looking at student work*. Workgroups might frame this as *aligning grading standards* and pose the question: Do we all agree what good work looks like? Alternatively, teachers might frame the same activity as *making sense of students' thinking about a concept*, posing the question: What different ways do students understand this idea? The first frame might support a discussion about scoring criteria, while the second might lend itself to identifying prevalent misconceptions. Both frames stand to support professional learning through the "same" activity but foreground different issues.

Data reports can be used within different frames, supporting different learning opportunities. In reviewing the effects of accountability systems on teachers' data use, Jennings (2012) identified five frames for this work: (1) *data as a lens*, or how educators view their schools, students, and themselves ("we are a failing school," "we are a high performing school"); (2) *data as a diagnostic tool*, or how educators determine what's working, what's going wrong, and why; (3) *data as a compass*, or how educators

decide what they should do in response to information uncovered; (4) *data as a monitoring device*, or how educators establish and monitor whether it worked; and (5) *data as a legitimizer*, or how educators justify decisions to themselves or to others. The different frames highlight how activities and the related epistemic stances conversationally position data reports.

Summary

The conversational features in Figure 1 work together in dynamic ways to support teachers' sensemaking in the service of our object of interest: professional learning opportunities. We see their coordinated use as manifesting different data use logics—instructional improvement or instructional management. By providing conceptual resources and orienting teachers to future work, these features allow a systematic comparison of learning opportunities in different workgroups during data use activities.

Research Design and Methods

Context of the Study

Our primary research question is: How do middle school mathematics teachers' data use conversations shape professional learning opportunities? This question arose during the course of our work on a larger project investigating instructional improvement in middle school mathematics in urban school districts. Since 2007, the Middle School Mathematics in the Institutional Setting of Teaching (MIST) project has investigated large-scale support of mathematics teachers' development of ambitious and equitable instruction.[1] Our research team identified urban school districts investing in viable strategies to meet these goals. The project examines mechanisms for instructional improvement from the district central office to schools and classrooms, collecting a variety of qualitative and quantitative data to document this work. We chose District B as one of our partners because of its investments in high-quality mathematics curriculum and intensive teacher professional development.[2]

District B is a mid-sized district, serving approximately 80,000 diverse students, of whom 50% identified as Latino, 25% African American, and 15% European American.[3] Over 25% of all students are classified as English language learners (ELLs). The majority of the students qualify for free or reduced-price lunch. District B's student achievement patterns in middle school mathematics look similar to those of many urban districts. On the annual state assessment for the year of this study, 50% of the African American students met mathematics standards, compared to 70% of the Latino students and 80% of the European American students. About 50% of the ELL students met standards in mathematics on these tests.

As a part of its efforts to improve mathematics achievement, District B invested in developing an interim assessment instrument and a curriculum framework, as well as providing teachers with collaborative time. The assessment instrument, which we call the ABC, was linked to the curriculum framework. Written by mathematics specialists in District B's central office, this document specified which standards should be taught when, accompanied by a pacing guide. The ABC provided accountability for the curricular framework by evaluating students on standards that corresponded with the previous 6 weeks' instruction. Mathematics coaches and teachers wrote the test items on the ABC over the course of several weekends. As a tool, the ABC assessments did not go through the same development process as the state test, lacking both field-testing and psychometric validation, yet were highly consequential. District personnel monitored schools through ABC reports. To assist in schools' own use of ABC data, the district provided teacher collaborative time as a chance to use ABC results to inform their instruction.

Data Collection

Data collection aimed to support a close analysis of teachers' data use conversations while also sufficiently capturing the broader context of the teacher workgroups. To serve the first goal, we sampled four meetings from each workgroup across the 2011–2012 academic year. To serve the second goal, we drew heavily on data from the MIST project. Because MIST's system-level questions required broad data collection, members of the research team distributed this work. The first and third authors spent time with participants and at meetings in the schools in this analysis, but a number of team members had a hand in recording meetings, documenting classroom instruction, and interviewing participants over time. During the study year, team members made 10 visits to the two sites described here.

Secondary data for the comparative analysis presented here include annual structured interviews with teachers, school leaders, and district personnel; annual surveys of teachers, school leaders, and other school staff; annual profiles of each focal school site's approach to instructional improvement in mathematics based on aggregated data and site visits; and a narrative summary of District B's evolving approach to performance data over the course of our study. In addition, we incorporate one of the project's quantitative measures, the *visions of high quality mathematics instruction* instrument (VHQMI; Munter, 2014). The VHQMI instrument provides a way to classify participants' views about the nature of good mathematics teaching around issues like the teachers' role, instructional discourse, and mathematics tasks.

Site and Participant Selection

To address our research question, we investigated how teachers used the ABC data to inform their instruction. While we did not have the capacity to closely examine all of the district's teacher workgroups to understand their data use practices, we followed the best-case logic of the larger study and sought out well-functioning workgroups, using an internal sampling technique (Bogdan & Biklen, 1992). That is, we asked key informants in the district to nominate teacher teams who collaborated well. We further refined our selection by interviewing participants and visiting sites. In the end, we identified 4 teacher workgroups in District B.

Focal Case Selection

Data use conversations were the primary unit of analysis. After identifying these in the recorded meetings, we selected two conversations for comparison because their similarities and differences stood to productively inform an analysis of teachers' learning opportunities. These conversations happened among two 7th grade mathematics teacher workgroups at Creekside and Park Falls Middle Schools, who used data from the same ABC test. The differences in data use proved informative in how specific practices shaped teachers' learning opportunities.

The schools' student and teacher demographics reflected those of the district, with a diverse group of students served by primarily European American teachers. Creekside's students were roughly split in thirds among students identified as African American, European American, and Latino. About 60% of the students qualified for free or reduced-price lunch. Among the mathematics teachers, about 65% identified as European American, with the remaining identifying as African American, Latino, or Pacific Islander. The 7th grade mathematics team averaged 5.3 years of experience (n = 8, SD = 7.6), more than the department average of 4.4 years (n = 12, SD = 6.8). The focal meeting included six 7th grade mathematics teachers; a special education assistant; a mathematics instructional coach, Tiffany; and the principal, Mr. Russell, who led the meeting.

Park Falls' students were identified as Latino (50%), African American (30%), and European American (10%), and about 75% qualified for free or reduced-price lunch. Eight of the school's nine mathematics teachers self-identified as European American. Overall, the teachers were less experienced than those at Creekside, with an average of 2.4 years of experience (n = 9, SD = 1.3). The 7th grade teachers had an average of 2.0 years of experience (n = 3, SD = .8). All three 7th grade teachers attended the focal meeting: Dennis, Claire, and Vanessa.

Data Analysis

We approached our comparative case study (Yin, 2009) with an interpretive stance informed by precepts of ethnomethodology (Garfinkel, 1967); that is, we sought to understand the meaning of participants' actions within their contexts. Unlike traditional ethnographic work, however, we did not have ongoing, everyday relationships with participants. Instead, we had to strategically mine MIST data to construct aspects of the cases. In this way, we offer a partially mixed-methods study, with an emphasis on the case-oriented qualitative research logic—what Leech and Onwuegbuzie (2009) describe as a QUAL + quan design. That is, qualitative logic dominates our inquiry, but it is augmented by some quantitative measures that capture aspects of the phenomenon under investigation and allow for direct comparisons.

Analyzing Focal Data

The focal meetings took place on 2 consecutive days in January 2012, about midway through the school year. We coded the meetings using constructs from our conceptual framework (Figure 1) to understand how the different conversational features shaped the workgroups' talk and, in turn, teachers' learning opportunities. As needed, we re-transcribed segments of interaction to conduct fine-grained discourse analysis to look at talk-in-interaction, adapting techniques from the ethnography of communication to support interpretations of meaning making (Saville-Troike, 2008). These additional markings capture tone, pace, and other metalinguistic details.[4] Heavily annotated transcripts trade in readability for the preservation of interactional detail, so we present the marked transcripts only as needed to support our analysis.

Describing Workgroup Cultures

To understand Creekside and Park Falls' workgroup cultures, we augmented the video analysis with other MIST data, especially the annual school profiles, focusing primarily on the 2011–2012 school year. We read the profiles for evidence of workplace norms. Since the MIST tracked mathematics instruction over time, the team developed and utilized a number of measures to gauge both teaching practice and discourses about mathematics teaching. The VHQMI involves using a set of interview questions that asks teachers to describe what they see as high-quality mathematics instruction and then coding responses according to a carefully developed rubric. Although the larger project quantifies these data to support large-scale analysis, because of our small sample, we report VHQMI qualitatively to understand influential case study participants' views on good mathematics teaching, an important component of individual epistemic stance. Because participants were

a subset of their school's mathematics teachers, we did not always have complete data on focal participants due to the larger project's design.

Describing Activity Structures

We looked across the eight sampled meetings, alongside participants' descriptions of typical meetings, to describe the activities and topics that usually occupied the workgroups' time. We inductively coded meetings by naming each activity segment (e.g., planning, data use, logistics). In this way, we could calculate how the workgroups allocated time within each meeting and across the sample. Percentages were obtained by direct calculation of time allocated, as represented in the videos ([minutes in activity / total meeting minutes] $\times$ 100). We compared the distribution of activities in the sampled meetings to participants' reports of how they used workgroup time. This coding also allowed us to revisit all instances of data use in the sampled meetings to compare the focal meeting to other instances of this activity. In this way, the closely analyzed conversations could be understood in relation to typical workgroup activities.

Understanding Resources for Professional Learning

The heart of this analysis uncovers the conceptual resources in teachers' data use conversations, with an eye on how these positioned teachers engage in future work. We divided each meeting into episodes, with boundaries determined by shifts in either topic or activity. From there, we looked at how the conversational features in Figure 1 were organized within each episode. Specifically, we identified problem frames invoked to understand interpretations of the topics at hand. We also looked at numerous representations of practice, including replays and rehearsals (Horn, 2005) and the representational forms of the ABC data reports. Additionally, we coded epistemic claims by marking any bald statements of what could be known or what was worth knowing through the data conversations. Together, these conversational features constituted the conceptual resources developed and deployed in the workgroups' conversations, reflecting each workgroup's prevailing data use logic. In addition to analyzing conceptual resources developed in conversations, we also attended to the second dimension of learning opportunities by noting how talk oriented teachers to future work.

Comparing learning opportunities in the two workgroups' meetings. In the final phase of analysis, we juxtaposed the two groups' conversational features and their contribution to teacher sensemaking, thus sharpening our understanding of the ways different data use practices mobilized teachers' future work and provided conceptual tools to support it. We summarized these professional learning opportunities in terms of their relationship to

broader accountability logics to illuminate the microprocess of teachers' data use (Little, 2012). We refined these claims by holding up the analysis to our secondary data, particularly the data use conversations in non-focal meetings and teacher interviews about the workgroups and data use practices.

Findings

Despite similarities in their teaching contexts and working from the same ABC assessment, the two workgroups engaged different data use logics, thus shaping qualitatively different teacher learning opportunities. In this section, we illustrate how an instructional management logic governed Creekside's data conversation, supporting learning opportunities that did not emphasize instructional change. Park Falls, on the other hand, engaged in both instructional management and instructional improvement data use logics, prompting more (but still limited) focus on instructional change. In the end, neither group was positioned to substantively improve instruction, undercutting a major premise of evidence-based practice in the current accountability policy.

Creekside's Instructional Management Logic: Increasing Proficiency Rates Through Resource Allocation and Individualization

We present our findings about the Creekside workgroup's data use. After describing the workgroup culture, its typical activities, and prevailing epistemic stance, we concentrate on the focal meeting, presenting the data use activity through our analytic framework to highlight teachers' sensemaking. We conclude with a summary of the meeting and the particular learning opportunities that the data conversation afforded teacher participants.

Workgroup Culture

A strong administrative charge to improve on AYP proficiency targets shaped the norms and values of Creekside's 7th grade workgroup. Having worked as Creekside's principal for 5 years, Mr. Russell attended the workgroup's weekly meetings and often emphasized this goal.

The accountability press extended beyond the workgroup meetings and into the broader school environment, from the introduction of all-day Saturday Math Camps to Mr. Russell's morning announcements reminding students to have a "good breakfast" on testing days. Creekside employed a full-time data manager to produce and monitor reports of both the ABC and state test. Representations of data and students' progress were displayed throughout the school, including on classroom walls labeled with children's names. The reports broadcast their stories everywhere, underscoring a strong orientation toward their narratives of student learning.

The teachers described Mr. Russell as highly involved in their classroom instruction. In interviews, they all reported that he observed them during the school year, with the majority of these visits being informal. Formal observations were accompanied by written feedback, creating a sense of instructional accountability. Mr. Russell worked closely with the school's mathematics coach, Tiffany. Together, they reported identifying mathematics teachers in need of support, and these people became central to Coach Tiffany's classroom-level work. As we found with other coaches in the MIST project, Coach Tiffany did not work with all mathematics teachers, due to their own resistance, perceived need, or her limited time.

Activity Structure

Across the sampled meetings, both Mr. Russell and Coach Tiffany facilitated workgroup discussions. These lasted about 35 minutes and took place during the school day. Teachers reported that they usually focused on planning and reflecting on lessons, sharing resources, and looking at ABC data to talk about improving scores.

Time allocation in the four sampled meetings reflected less planning and more administrative tasks. Looking at achievement data was a primary activity in the sampled meetings, accounting for 27% of the time, followed by an entire meeting devoted to the school's Improvement Plan (26%), lesson planning (17%), and planning for tutoring sessions (8%).

Epistemic Stance

Aside from the ways epistemic stance was communicated in interaction, we saw the prevalence of certain ideas about mathematics, teaching, and learning through the VHQMI measure. Using VHQMI to classify participants' ideas about good mathematics teaching, we found that Coach Tiffany had a vision of mathematics teaching that was more ambitious than the majority of Creekside's teachers, signaling greater engagement with instructional ideas aligned with documents like the National Council of Teachers of Mathematics' (2000) *Principles and Standards*. Mr. Russell, however, had a less ambitious vision of good mathematics instruction than Coach Tiffany and many of the teachers. Given his leadership role, this may have played a part in the *instructional management* emphasis of the meetings, since his ideas about good mathematics instruction were less developed.

Focal Episode: Conversational Sensemaking in Data Use Activities

With the recent completion of the third ABC assessment of the school year, the January 27, 2012, meeting focused on teachers' data analysis. Mr. Russell facilitated the majority of the meeting. In preparation, he had asked the school's data manager to produce tables listing the school's

Figure 2. **Creekside teachers looking at state test and ABC proficiency-coded data arrays for their African American students.**

African American students. The tables included the prior year's state test scores alongside scores on the three ABC tests from the current year. His focus on African American students came out of NCLB's design, which holds schools accountable to the performance of subpopulations. At Creekside, the school's AYP targets included measurable improvement among these students. This data representation, with its focus on one student subgroup, communicated a stance on what mattered for these teachers' future work. The scores were color coded to show each student's proficiency level on the assessment, using the categories *commended, passing, bubble*, or *growth*. The data manager had disaggregated the tables by teacher so individual teachers could look at their own students (see Figure 2). The representation brought together particular students and the labels related to the accountability system, placing those two frames for interpreting teaching in simultaneous focus.

Mr. Russell's agenda and time allocation also contributed to the activity's meaning, making certain aspects of instruction salient while keeping others out of view. The activity was divided into three segments: an introduction by Coach Tiffany reviewing new standardized testing procedures (3 minutes), data use (27 minutes), and looking ahead to the next unit of instruction and the next meeting (8 minutes). We focus primarily on the meanings developed through the teachers' data use conversation.

Mr. Russell's 5-minute overview of the data use activity invoked both the compass and monitoring frames (Jennings, 2012). Recall that the compass frame orients conversation toward questions about what to do in response to the information that data reports uncover, while the monitoring frame uses data to establish whether or not previous decisions are working. As he explained, the purpose of the activity was to use the raw scores to predict students' proficiency on the state test (*compass*), saying,

> The importance of this, before we get started, is to dig deep to make sure that we're meeting kids' individual need, we're talking to them about their test scores, that we're having those conversations that motivate them.

The particulars of Mr. Russell's compass framing communicated an epistemic stance that effective teaching involves individualization and motivation, orienting teachers to these values. He also represented good instruction, specifying the importance of "meeting kids' individual need" and "having those conversations that motivate them." Together, the frame and the representation orient teachers toward images of desirable future work.

Later in his introduction, Mr. Russell also invoked a monitoring frame (*how are we doing?*):

> We're gonna see if what my [predicted pass rate] numbers look like when I go through and do this based on their data alone matches up with your numbers. And you have much more data 'cause you've got the perception data in place. You know the kid better than I do, obviously.

Along with the monitoring frame, Mr. Russell inserted an epistemic claim that the teachers had "the perception data in place," suggesting that their knowledge could enhance the data and support inferences about student learning. Invoking that epistemic stance also shaped teachers' engagement in data use, as their (presumed) greater knowledge of students mattered.

The teachers continued primarily in the compass and monitoring frames that Mr. Russell introduced, spending the next 15 minutes poring over the data and annotating the tables to predict student performance on the state test (*compass*). Teachers worked individually without much discussion, but comments caught on video included a number of epistemic claims. These ranged from linking test performance to resources (e.g., "Tutoring is helping some of them") to questioning how well the test reflected student knowledge (e.g., "Jaylen has test anxiety; he probably stresses out"). These examples fell in with the compass and monitoring frames, and other conversational features augmented them to support sensemaking. In the first example, "tutoring is helping" served as an epistemic claim about their collective work, while the second represented Jaylen in a way that supported a modified interpretation of his individual data.

Only one recorded epistemic claim invoked a diagnostic frame (*What's working? What's going wrong?*), as a teacher tried to make sense of students' learning ("They understand the process, but applying it, that's hard"). Notably, the grounds for this epistemic claim in the diagnostic frame were indeterminate in our records. It did not come from the reports themselves, which, as representations, showed only summative scores without specifying questions or topics.

In the next segment of the activity, Mr. Russell asked the teachers to go around and report their predicted numbers while he recorded them. Teachers presented numbers as lists without much commentary (e.g., "I have 18% passing, 27% bubble, 55% growth"), with only the occasional clarifying question interjected (e.g., "No commended?"), returning the activity to a monitoring frame. As an activity, the minimal dialogue around the newly formulated numbers communicated that they adequately represented student learning.

In the end, Mr. Russell reported the findings of the exercise, saying,

> When I broke down the numbers, just based upon data, because I don't get to see the kids every day, I was sitting somewhere around a 55 to 58% of the kids we could go get and somewhere in the neighborhood of 40 to 43% of the kids were just working on growth, which is kind of alarming numbers. . . . You guys think more kids you can get over the hump, and that's encouraging. . . . So, you guys are saying, you know, that number's probably closer to 60, 65% of kids that we can go get either, you know, bubble, passing, or commended kids and that's encouraging news, especially with [the] most difficult sub-pop.

The claim that closed out this turn of talk ("You guys think more kids you can get over the hump") both implicated the teachers' potential and also held them responsible for seeing their predictions through.

The ensuing 2-minute discussion focused on how teachers could inculcate that growth in the target students (*compass frame*), further specifying the next steps they should take. Teachers discussed sending students to Math Camp, getting them to tutorials, personalization, motivation, and paying attention to individuals' engagement in class, further detailing the nature of their future work. Coach Tiffany extended these images, linking them back to the classroom:

> I think it's easy to overlook the kids that're borderline, especially if they're real quiet in class. We sometimes overlook them. And so, just to make us open our eyes and say, "Okay, you know, this student, they're, they keep to themselves and they participate. They're still a bubble kid and they're still a child that could be either pushed up to commended or passing." So, just kinda open our eyes as to what kids are in our room, just a reminder.

In this representation of future work, Coach Tiffany rehearsed how teachers could think about quiet, borderline students who might be overlooked. By including herself in this image with the first person plural ("we sometimes overlook them"), she normalized this problem (Horn & Little, 2010) as a collective issue while mobilizing their future work ("kinda open our eyes"). She also highlighted specific practices the teachers could use. While attending to classroom inclusion may improve an aspect of teaching, we do not see this as supporting deeper instructional change.

The personalization and motivation discussions provided the teachers with conceptual guides. Together, the conversational features and images of future work shaped their learning opportunities. Many teachers' responses indicated their enjoyment of the activity, suggesting that they assented to this future work. In fact, they said they hoped to get these kinds of data on all their students, not just the African American "sub-pop." Analytically, we view this as an endorsement of the desirable future practice set up through this conversation.

Mr. Russell explained the work entailed in producing the tables ("[The data manager] has to do it manually") and was non-committal about making the data available for other students in this way ("We can try. . . . I won't say no"). Resource allocation came up as a general concern, capping possible responses that the compass frame (*what are sensible next steps*) might imply. Specifically, Mr. Russell said,

> We have to worry about resource allocation; let's make sure we're getting the right kid in Camp that's gonna behave correctly. I mean, I know that's difficult, but definitely something we have to focus on at a large school with limited resources.

Mr. Russell's closing statement limited the possibility of extending this data activity for other groups of students, all but foreclosing making this routine in the future. More subtly, his statement put a threshold on which bubble kids merited scarce resources, charging the teachers to "make sure [they're] getting the right kid in Camp that's gonna behave correctly." Although the earlier indications of the teachers' future work invoked an inclusive vision of instruction, the exclusion of poorly behaved students permitted teachers to limit the targets of their attention.

Conversational Learning Opportunities: Data Use Logics at Creekside

The Creekside teachers' conversations rendered certain aspects of teaching meaningful while leaving others less visible. The meeting mobilized teachers for the future work of individualizing and personalizing instruction for students on the cusp of proficiency ("bubble kids"), naturalizing accountability categories in teachers' sensemaking, signaling this meeting as an opportunity to learn about which students needed extra resources to meet proficiency targets.

Because of its emphasis on proficiency targets via resource allocation (including Math Camp, tutorials, and teacher attention) for *particular* students, this conversation was primarily governed by the instructional management logic. The prospective work of engaging students predominantly addressed the problem of improving test scores without substantively re-thinking the work of teaching, thus providing teachers with learning opportunities about redirecting their attention—and very little about the instructional nature of

that attention. Although interactions rendered fleeting images of teaching, they largely attuned teachers to particular students in their classrooms ("the quiet kids") rather than mobilizing them to change instruction itself.

Thus, the learning opportunities organized by this conversation, driven by instructional management logic, gave few inroads to rethinking the particulars of mathematics teaching. The summative data scores simply represented whether students had passed: they did not point to troublesome topics, leaving the details of students' mathematical learning to teachers' "perception data." By relying on teachers' impressions to fill in these details, the evidence basis of this conversation simply augmented the typical conversations that teachers would have absent the test reports. By excluding critical issues of mathematical learning, the majority of the conversation avoided some of the potentially richest source of supporting African American bubble kids—and all students. In other words, Creekside teachers had few opportunities to learn *how* to individualize instruction for students, but ample opportunity to learn *whom* among a certain subset of students needed this individualization and motivation. Finally, there was little attention to the underlying reasons that African American students might be lagging in achievement scores or what it might mean for the mostly white teachers to build motivating rapport, marking this as a colorblind conversation (Yoon, 2012). Notably, the 8-minute planning session at the end was divorced from the data use conversation.

Park Falls' Instructional Improvement and Management Logics: Interpreting Students' Mathematical Understanding

We present our findings about the Park Falls workgroup's data use. After describing the workgroup culture, its typical activities, and its prevailing epistemic stance, we concentrate on the focal meeting. We provide an overview of the data conversation, describe an emergent interactional routine for looking at data, and present a close analysis of one instance of that routine to illustrate the conversational sensemaking at work. We conclude with a summary of the meeting and the learning opportunities that it afforded participants.

Workgroup Culture

Our project-level profiles of Park Falls described a school work environment that, on the whole, was unstable and conflicted. Despite this, the 7th grade mathematics teachers' meetings revealed a workgroup that was highly committed to their students. Park Falls had been reconstituted during the 2010–2011 academic year, getting a new principal whose mandate was to "shake things up." Right before the 2011–2012 school year started, however, this new principal resigned and a third one was assigned to the school on the second day of classes. In interviews, the mathematics teachers reported that the new principal was not an ongoing presence in their classrooms,

infrequently observing formally or informally. The teachers also reported that the school-level mathematics coach was not helpful, leaving them largely on their own when it came to instructional improvement. Additionally, interviews with school participants uncovered tension between the coach and principal. The coach herself reported frustration in sorting through conflicting demands from the district and principal, with the former asking her to support teachers and the latter asking her to analyze ABC data.

Activity Structure

The 7th grade team met twice a week for 40 minutes. The coach would occasionally lead the meetings, and while the principal occasionally attended, she did not lead. The principal expressed interest in having teachers use ABC data, although teachers did not report data use as a central activity in their workgroup. The teachers told us that the workgroup meetings were used to lesson plan, discuss struggling students, compare classes, and look at student work together. Our sample reflected much more data use. Across the sampled meetings, time was allotted to data use (47%), lesson planning (25%), designing assessments (16%), and administrative or logistical discussions (11%), making the focal meeting representative of the group's activity.

Epistemic Stance

As with Creekside, we used the VHQMI instrument to uncover the dominant views on the nature of good mathematics teaching among the Park Falls' staff. The mathematics teachers' view of good teaching was, on the whole, traditional, as were those of the coach and principal. We did not have the VHQMI data on all the 7th grade teachers, but Claire stood out for her more sophisticated vision of good mathematics teaching on the VHQMI scales.

Constructing Meaning in Data Conversations

In the January 26, 2012, meeting, the 7th grade teachers, Dennis, Claire, and Vanessa, gathered to look at the latest ABC data. There was no formal facilitator, but two of the three teachers had prepared for the meeting by reviewing the test. Claire had led a discussion with her students about frequently missed items to better understand their confusion, and Dennis took the test himself to understand what the questions were asking of students mathematically.

The teachers' activity invoked a diagnostic frame as they reviewed and interpreted frequently missed items from the ABC test. As a guide, they collated a number of representations: the test booklet, a printout showing the distribution of student answers on each multiple-choice item, and Dennis's copy of the worked test (see Figure 3).

Over the course of the meeting, a conversational routine emerged (Horn & Little, 2010). First, the teachers used the item response distribution to identify

Figure 3. **Park Falls teachers' representational array included the worked ABC test, blank test booklets, and the distribution of student answers on each item.**

a frequently missed question. Then, they looked at the distribution of responses to discern what students might have been doing, attending to high response rates to distractors. As they interpreted students' responses, Dennis and Claire drew on their preparation for the meeting, with Dennis describing taking the test and Claire reporting what students said about their thinking. They would then go to the next frequently missed item. In the end, they reviewed 18 of the 30 test questions using variations on this routine, spending on average just under 2 minutes per question. During their conversation, they generated analyses about student responses that ranged from why students skipped questions, why notation was confusing, how students mixed up similar sounding words, and how students' limited understanding was reflected.

The tacit assumption seemed to be that understanding student thinking would support more effective instruction, which aligns well with an instructional improvement logic for data use. Notably, the details of future instructional activity, while more closely organized around students' thinking than they were at Creekside, remained underspecified.

A Closer Look at Conversational Sensemaking

To elucidate the learning opportunities supported in the Park Falls workgroup's conversation, we present a close analysis of one instance of the emergent conversational routine. This example illustrates how the conversational features in Figure 1 worked in concert to organize the teachers' learning opportunities. Our exposition of this conversation is not parallel to the exposition of the Creekside teacher workgroup. The rich interactions here warranted the use of fine-grained transcription and analysis.

Figure 4. **Screen shots of the closely analyzed episode from Park Falls. The teachers coordinated meaning across representations by using gesture and talk to draw collective attention to different information in the data reports. The numbers indicate where the gestures correspond to those coded in the transcript.**

The following excerpt began almost 12 minutes into the meeting and lasted for 2 minutes and 23 seconds, making it slightly longer than some of the other item analysis episodes. Figure 4 provides screen shots to show how teachers coordinated the data reports and the test questions to interpret student responses. We also present the transcript, marked for metalinguistic qualities of talk, to give a sense of participants' extensive involvement in this conversation (Horn, 2010): the ways they completed each other's sentences, inserted acknowledgment tokens ("mhm"), and overlapped speech signaling collaborative sensemaking. The right hand column of the transcript includes gestures that accentuated the speakers' meanings. They are noted with numbers to help the reader coordinate talk and gesture.

This episode began when Claire pointed to Question 7 in the test booklet, the next frequently missed item on the test, which read,

> The diagram shows a floor plan of a house. Each 1 centimeter on the diagram equals 4 meters on the actual house. What are the dimensions of the master bedroom in the actual house?

A floor plan diagram, shown in Figure 4, accompanied the question. In it, a rectangle labeled "master bedroom" contained the numbers 1.5 × 2. Students were given four possible responses:

(A) 1.5 × 2 m (B) 4 × 4 m (C) 6 × 8 m (D) 10 × 4 m.

The correct answer is C, since 1.5 (the width) × 4 (the scale factor) = 6 and 2 (the length) × 4 (the scale factor) = 8.

The frames invoked during the discussion of this question were primarily diagnostic (*what is going wrong and why?*), with some compass (*what should they do in response?*).

1	CLAIRE:	This one over here? *(1)* Seven? It, it messed with them because they didn't understand that (.) 1.5 and 2? *(2)* like I had kids try to measure that *(3)* and they're like, "But it's not 1.5. It's not 2," and I was like, *(4)* "They're giving you your dimensions." *((laughs))*	*(1) Tapping on test booklet, pointing to question* *(2) Pointing to master bedroom rectangle on top left* *(3) Indicates length of rectangle with thumbs* *(4) Throws hands up over rectangle*
2	VANESSA:	Yeah, they didn't understand *(5)* and I think it would've helped, too, if the actual question *had* centimeters on there. Even though I *know* (.) that this is obviously a floor plan=	*(5) Waves pen over same problem on Claire's booklet*
	CLAIRE:	=Right=	
4	VANESSA:	=This is not the actual dimensions of a house.	
5	CLAIRE:	Right.	
6	VANESSA:	And I think it would've helped if it said 1.5 centimeters by 2 centimeters because then, you would realize, "Oh, that's=	
7	CLAIRE:	=Yeah.=	
8	VANESSA:	=because=	
9	CLAIRE:	=I=	
10	VANESSA:	=it's written here and they read that.	

Using the representation of the test booklet, the teachers looked at the diagram to interpret their students' confusion. Claire noted that students did

not understand the "1.5 × 2," pointing to the diagram that showed that the master bedroom was labeled with dimensions, but that the dimensions were not placed along their length and width (see Figure 4). In turn 1, Claire replayed what students did during the testing event, sharing her observation of students attempting direct measurement of the floor plan ("I had kids try to measure that"), suggesting that they understood something of *length* and *width* but not the way they were shown in the diagram. Vanessa endorsed and extended Claire's interpretation of the students' confusion in turns 2 to 10, adding that it would have been helpful to have units on the numbers to help students understand that they labeled the rooms' dimensions. Here, the teachers critically evaluated the test item while making sense of their students' understandings of the mathematical content.

A few turns of talk later, Dennis built on Vanessa's observation about the absence of units as a source of confusion.

13	DENNIS:	//It was] really tacky that they didn't put units? on this question. Just=	
14	VANESSA:	=Well=	
15	DENNIS:	=I, I stressed that *so* much in class, "Put your units so there's *no* confusion"=	
16	VANESSA & CLAIRE:	=Right.=	
17	DENNIS:	//And there're no]	
18	CLAIRE:	//Especially, even when they set it up] Like	
19	DENNIS:	It, it helped 'em *see* that it causes confusion when there's no units=	
20	VANESSA:	=Uh huh=	
21	DENNIS:	Like I think they had, they learned that (.) it's good. But for me, I had to sit there and think through:: (.) through it. And there should be something on here also that says it's not to scale.	
22	VANESSA:	(.) Right= *(7)*	*(7) Tapping distribution sheet to row for this question*
23	CLAIRE:	=Yeah.	
24	DENNIS:	Like, because that was:: =	
25	VANESSA:	=I agree=	
26	DENNIS:	=another:: (.) issue I had here was it's not (.) to scale.	

Dennis offered two observations about the item design to extend Vanessa's critique. Although he did not explicitly say that either of these contributed to student confusion, this was perhaps implicit in the conversational

context. First, he replayed his own instructional emphasis on the importance of using units to make numbers meaningful (turn 15: "Put your units so there's *no* confusion"), contrasting it with how numbers were marked on the test item. In turn 21, Dennis further noted that the diagram's lack of scale made it unclear and was not communicated, supporting this by reporting his own experience taking the test.

In the next segment of talk, Vanessa brought another data representation into the conversation. She looked at the distribution of student responses to Question 7 to understand how widespread the students' errors might have been.

27	VANESSA:	*(8)* One out of every four kids put A on the question.	*(8) Pointing to distribution sheet*
28	CLAIRE:	Yeah, they just	
29	VANESSA:	//It's incorrect] *(9)*	*(9) Taps on A on the booklet*
30	CLAIRE:	//picked it.]	
31	DENNIS:	//Oh yeah.]	
32	CLAIRE:	//They just picked it.]	
33	DENNIS:	That's it?	
34	VANESSA:	*(10)* Yeah. Exactly what that said.	*(10) Points to distribution sheet*
35	DENNIS:	Only that, I figured it'd be way more.	
36	VANESSA:	No, not that *(11)*. Only 60% got it right though.	*(11) Taps cell with the data on it*

Here, the teachers coordinated two other representations: the distribution of students' correct and incorrect responses and the content of the responses themselves. This coordination stood to support the diagnostic frame of the activity as they further made sense of student thinking. Notably, answer A required students to correctly identify the master bedroom on the diagram but it did not require the proportional reasoning that the test item intended to assess. Perhaps in line with their earlier observation about unlabeled units, Dennis was surprised that "A" was not a more frequent response (turn 35), since the unlabeled numbers were a potential source of confusion. However, the teachers did not draw more inferences about students' thinking from the distribution of student responses, perhaps because one out of four responses would be consistent with guessing, an observation they made in discussing another item.

Nonetheless, the teachers turned their attention to instructional implications, invoking a compass frame.

37	CLAIRE:	Um. (.) And I think one thing. That threw them (.) on it was that (.) like they're used to seeing it (.) like this? *(12)* like on sides of things?	*(12) Makes gesture outside camera frame*
38	VANESSA:	Yeah.	
39	CLAIRE:	And now it's (.) it, *(13)* this is telling you your di*men*sions *(14)* but like, even on our homework this week, there's something asking 'em, "What's the dimensions of this (.) volume?"	*(13) Points to master bedroom rectangle on top left of question in test booklet* *(14) Taps rectangle*
40	VANESSA:	That's//	
41	CLAIRE:	And they don't know what *dimensions* are. They don't understand that this *(15)* is like a 2 by 4 or whatever, like 2 *(16)* by one and a half *(17)*.	*(15) Points to dimension label in master bedroom rectangle* *(16) Runs finger across length of diagram* *(17) Runs finger across width of diagram*
42	VANESSA:	Yeah.	
43	CLAIRE:	And so, that *(18)* threw them because it wasn't on the sides.	*(18) Taps the diagram*
44	VANESSA:	Yeah, I had kids ask me *(19)* what dimensions *were* (.) for the homework, and //I said]	*(19) Taps diagram with pen*
45	CLAIRE:	//Me too]	
46	VANESSA:	And I said, "So, what would be the di*men*sions of this classroom? What would it be here?" And they're like, "Oh, yeah, length and," they knew, but they don't think of it as (.) di*men*sions.	
47	CLAIRE:	Mm-mm.	
48	VANESSA:	So, that word is, maybe we should=	
49	DENNIS:	=Well=	
50	VANESSA:	= // hit that word]	
51	DENNIS:	// it's one] of those words that has multiple meanings=	
52	CLAIRE:	=Yeah. //That's true.]	
53	VANESSA:	//Righ::t]	

This segment of talk began to flesh out the teachers' future work in relation to their analysis of this item. In turn 37, Claire repeated Dennis's earlier claim that students were accustomed to seeing dimensions labeled. She supported her observation by representing a recent classroom observation. In turn 39, she replayed a homework question that asked about dimensions, making an epistemic claim in turn 41 that "they [the students] don't know what dimensions are." Vanessa confirmed her observation, representing a similar observation in her class (turn 44) and then replaying her instructional response in turn 46 ("what would be the dimensions of this

classroom?") as a way of helping students develop their understanding of *dimension*. The teachers ended by collectively deciding to help students better understand the idea of *dimension* since *dimensions* is "one of those words that has multiple meanings" (Dennis, turn 51). Although the instructional mathematical goal was clear, the specific course of action for future teaching was not fully articulated ("maybe we should hit that word"; Vanessa, turn 50).

Conversational Learning Opportunities: Accountability Logics at Park Falls

The Park Falls teachers' conversation centered squarely on student thinking, with their analysis of frequently missed items and interpretations of student errors. This activity mobilized teachers to modify their instruction in response to identified confusion. Therefore, this conversation afforded learning opportunities about what topics need revisiting and why, supported by a data-based interpretation of students' thinking.

Comparing the closely analyzed episode to other instances of Park Falls' item discussions, we see this interaction as representative of the Park Falls teachers' data use practice. The representations available to teachers in this meeting—a distribution of student responses and the ABC test booklets, along with replays and rehearsals of instructional interactions—afforded the meeting's activity structures and problem frames. Supported by this representational array, teachers interrogated the data report's narrative, conversationally positioning it as open to inquiry and talking back to its findings, such as when they questioned the placement of the dimensions and noted the lack of unit labels as potential sources of confusion. The objective veneer of the ABC data report thus wore thin, as considerations about the relationship between students' mathematical thinking (e.g., do they understand dimension and scale?) and the particulars of the testing event (e.g., Claire's observation about the students' attempts to measure the diagram directly) could be put alongside classroom observations (e.g., Claire and Vanessa's replays of homework trouble on a dimension question) to interpret the resulting student performances and what they suggested about mathematical understanding.

Unlike the conversation at Creekside, then, this discussion uncovered many details of students' mathematical thinking, from their limited grasp of certain topics to miscues resulting from the test's format to misalignments with instruction. While the data report, positioned as one source of inference, supported intensively diagnostic conversation, the teachers did not spend as much time extrapolating next steps for instruction in a compass frame. In the closely examined episode, for instance, they agreed to "hit hard" on *dimension*, but the details of how to do so remained opaque. Instead, the teachers, guided by the report, moved on to the next frequently missed question. Thus, we see an instructional improvement logic at work in

Park Falls' data use—they know what to revisit and why—counterbalanced by an instructional management logic in pressing through the data report to ensure a comprehensive analysis in limited time.

Discussion

Both Creekside and Park Falls faced sanctions over students' performance on state tests, making data reports highly consequential narratives of teaching and learning in each school. We examined how each workgroup's data use informed instruction—what we could reasonably call evidence-based instructional practice. As Table 1 summarizes, Creekside and Park Falls operated under different data use logics, giving each conversation a distinctive character. These differences, manifested through differences in the conversational features, supported distinctive meaning making and, in turn, dissimilar learning opportunities.

The data use logics in these meetings reflect the workgroups' broader cultures. Creekside's instructional management logic aligned with the strong accountability press from the principal, Mr. Russell, who had a limited understanding of good mathematics instruction yet needed to improve his school's test scores. Mr. Russell had the school's data manager represent the student performance data to identify target African American students, who underperformed on standardized assessments, for extra personalization and instructional attention, without providing other data-based representations for discussing the nature of the students' mathematical understanding. In this way, his conception of the problem—which targeted a subpopulation needing improvement according to the AYP report, emphasizing personalization and motivation while glossing over the particulars of students' mathematical understanding—got built into the activity.

Monitoring and compass frames guided the Creekside teachers' data use, with almost no diagnostic frames to uncover *why* their students did not learn as expected. Because of the available representations, these details were not accessible. This example illustrates the interdependence of conversational frames and representations: it is hard to work in a diagnostic frame absent certain information, and the data representations limited what teachers could access. In the end, the conversation supported teachers' learning opportunities by organizing their future work to attend to issues of student motivation and the allocation of additional resources (e.g., who needs more personal attention, who needs to attend Math Camp). Both the nature of mathematics instruction and the particulars of building rapport were, for the most part, left out of the conversation, limiting the possibility of instructional improvement.

In contrast, the Park Falls workgroup's data use engaged both instructional improvement and instructional management logics. A diagnostic frame primarily shaped their activity, as the teachers sought to understand

Table 1
Summary of the Sensemaking Resources and Learning Opportunities at Creekside and Park Falls

	Creekside	Park Falls
Workplace culture	• Strong administrator • Press and accountability from principal	• Unstable leadership • Tensions between coach and principal
Activity structure	• Weekly grade level meetings • Identify bubble kids among African American students to target for intervention	• Twice weekly grade level meetings • Look at frequently missed items on interim assessment to figure out adjustments to instruction
Frames	• Compass and monitor frames • Categorizing students as commended, passing, bubble, or growth using "perception data" with test data	• Diagnostic and compass frames • Looking at frequently missed items and making sense of why students missed them
Representations	• Spreadsheets of African American students' scores, disaggregated by teacher	• Distribution of student scores • Test booklet • Replays of students' feedback
Epistemic claims (*examples*)	• "Meeting kids' individual needs" • "Having conversations that motivate them" • "It's easy to look over kids that are borderline"	• Students skip questions that are too long • The notations and diagrams confuse students • Students have limited understanding of mathematics
Learning opportunities	• *Conceptual resources:* Individual attention and strategic resource use will assist African American bubble kids • *Future work:* Providing attention and resources to identified students	• *Conceptual resources:* Analysis of why students missed test questions, with attention to the items, instruction, and student understanding • *Future work:* Re-teaching topics with attention to sources of confusion

what the data revealed about what was working and why, with some brief moments engaged in a compass frame, as they discussed possible

instructional responses. In line with this, their main data use activity focused on students' mathematical thinking and test-taking experiences, with the compass frame pointing their future work toward some adaptations in instruction. By coordinating their interpretations across an array of representations, the teachers could make strong inferences about students' mathematical thinking. While the representations of data limited the Creekside teachers' access to diagnostic frames, the press to review all of the frequently missed items limited the Park Falls teachers' compass frame: they simply did not have time to discuss instructional implications deeply. This tension between instructional improvement and management had consequences for teachers' learning opportunities, however. Although their conversations supported an analysis of student understanding, their future instruction remained underspecified. The knowledge may have been, in Mandinach and colleagues' (2008) terms, "usable," but the activity structure limited their capacity to use their findings with their students in the most thoughtful ways possible.

Conclusion

Injunctions for evidence-based practice cannot be understood apart from teachers' communities, whose norms, practices, and values shape the interpretation of evidence. This study offers a framework for identifying the logics that drive data use activities, along with specific conversational features that shape learning opportunities therein. This analysis contributes a microprocess view of teachers' data use and its relation to teachers' opportunities to learn about instruction, as well as a framework for leading and shaping those discussions in productive ways.

We conclude with two implications of our study, followed by an observed irony resulting from the current accountability culture. First, while we contend that the conversational features identified shaped teachers' learning opportunities, the quality of these learning opportunities did not come out of any single conversational feature. We could imagine a less diagnostically rich conversation among a different workgroup who used the Park Falls teachers' array of data representations. The interplay among the conversational features identified in Figure 1 worked in concert with overarching data use logics, buttressed by Claire's (and, we suspect, Dennis's) instructional expertise, to support the workgroup's sensemaking. Thus, particular forms of conversational features may be necessary but not sufficient for instructional improvement. Nonetheless, we suggest that teachers' data use can be productively characterized by this framework and realigned toward the goal of instructional improvement.

Our second, related point is that, in some instances, the particular forms of conversational features *limited* the workgroups' learning opportunities. At Creekside, the data representations reduced learning opportunities by

obscuring critical aspects of teaching and learning. Specifically, the absence of scores on specific mathematics topics, along with a focus on a subset of their students, narrowed the teachers' ability to make strong inferences about what their students knew and understood. At Park Falls, the teachers' activity was organized around the goal to review *all* the frequently missed items. We saw their conversation press forward at a quick pace (about 2 minutes per test item), leaving little time to specify how they might incorporate their thoughtful diagnostics in their future instruction. Without clear insights into the nature of students' understandings (Creekside) or adequate time to construct meaningful instructional responses (Park Falls), both groups' learning opportunities were limited.

These two points lead us to our observed irony: The limitations in teachers' learning opportunities identified in both workgroups emerged *specifically* in response to the accountability policy. NCLB's demands worked against instructional improvement in many ways. We see Creekside's narrow focus on allocating resources to (well-behaved) African American bubble kids and Park Falls' rush to address all test items as a direct result of the accountability policy's structures. Yet, these were precisely the aspects of their conversations that truncated deeper reflections on instruction. Our broader project data suggest that these policy-rooted structural limitations on learning are not unique to these workgroups.

Adding to this irony, we suspect that the Park Falls teachers' stronger pedagogical content knowledge may have hampered their uncritical use of the ABC data. Based on our own examination of the test items, we concur with many of their critiques. However, this critical stance led the teachers to discount the evidence to which they were nonetheless beholden in order to meet their school's proficiency targets. This put the teachers in a conflicted position: do they follow their own sense of students' understanding, supported by a thoughtful critique of the assessment tool, or do they meet the policy's demands and teach their students to become better test takers? This is obviously not an either/or proposition, but we saw the Park Falls teachers dismiss some of the report outright based on their judgments of the quality of the test items.

If instructional improvement is to be a meaningful goal of data use, trustworthy accounts of teaching and learning through well-designed assessment tools are vital. We suggest that states and districts develop transparent and meaningful mechanisms for teacher feedback around assessments to increase that trust. We see many calls for improving accountability as a mechanism for instructional improvement by investing in teacher training around data use—and very few for inviting teacher feedback on the assessments that generate those data. In the end, if assessments do not provide valid and reliable diagnostic information to support instructional improvement, the instructional management logic will, of necessity, overtake the whole standards-based accountability enterprise, because it will be all that matters.

Notes

The authors would like to thank the MIST research team, particularly Mollie Appelgate, Diana Baldys, Jason Brasel, Charlotte Dunlap, Erin Henrick, and Brooks Rosenquist. Special thanks to Paul Cobb and Kara Jackson for comments on earlier versions of this article. The opinions expressed here reflect the views of the authors. The research was supported by the National Science Foundation under Grant No. DRL-1119122 and the Spencer Foundation.

[1]Paul Cobb was the principal investigator (PI). Erin Henrick, Ilana Horn, Kara Jackson, and Thomas Smith were co-PIs.

[2]All proper names are pseudonyms.

[3]Numbers are rounded to prevent reverse-look ups that might identify our sites.

[4]Talk turns numbered for identified speakers. Continuous speech at turn boundaries shown with =equals signs, while onset of [overlapping talk is shown with left brackets and conclusion marked with a slash/. EMPHATIC talk shown in caps, and elong:::ated enunciation shown with repeated colons. ((*Activity descriptions*)) appear within double parentheses and in italics.

References

Bannister, N. A. (2015). Reframing practice: Teacher learning through interactions in a collaborative group. *Journal of the Learning Sciences*. DOI: 10.1080/10508406.2014.999196

Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education: An introduction to theory and methods* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas Accountability System. *American Educational Research Journal*, *42*(2), 231–268.

Diamond, J. B., & Cooper, K. (2007). The uses of testing data in urban elementary schools: Some lessons from Chicago. *Yearbook of the National Society for the Study of Education*, *106*(1), 241–263.

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

Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.

Goffman, E. (1974). *Frame analysis*. Boston, MA: Northeastern University Press.

Greeno, J., & Gresalfi, M. (2009). Opportunities to learn in practice and identity. In Future workP. A. Moss, D. C. Pullin, J. P. Gee, E. H. Haertel. & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn* (pp. 170–199). New York, NY: Cambridge University Press.

Hall, R., & Horn, I. S. (2012). Talk and conceptual change at work: Adequate representation and epistemic stance in a comparative analysis of statistical consulting and teacher workgroups. *Mind, Culture, and Activity*, *19*(2), 240–258.

Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices toward a policy and research agenda. *Educational Policy*, *22*(4), 578–608.

Horn, I. S. (2005). Learning on the job: A situated account of teacher learning in high school mathematics departments. *Cognition & Instruction*, *23*(2), 207–236.

Horn, I. S. (2007). Fast kids, slow kids, lazy kids: Framing the mismatch problem in mathematics teachers' conversations, *Journal of the Learning Sciences*, *16*(1), 37–79.

Horn, I. S. (2010). Teaching replays, teaching rehearsals, and re-visions of practice: Learning from colleagues in a mathematics teacher community. *Teachers College Record*, *112*(1), 225–259.

Horn, I. S., & Kane, B. D. (in press). Opportunities to learn in mathematics teachers' collaborative conversations. *Journal of the Learning Sciences*.

Horn, I. S., & Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions. *American Educational Research Journal*, *47*, 181–217.

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the "data-driven" mantra: Different conceptions of data-driven decision making. *Yearbook of the National Society for the Study of Education*, *106*(1), 105–131.

Jackson, K., Cobb, P., & Rigby, J. G. (2014, October). *Instructional improvement and instructional management: District leaders' orientations towards improving mathematics teaching and learning*. Paper presented at the University Council for Educational Administration, Washington, DC.

Jennings, J. (2012). The effects of accountability system design on teachers' use of test score data. *Teachers College Record*, *114*(11), 1–23.

Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education*, *112*(4), 496–520.

Latour, B. (2009). *Reassembling the social: An introduction to actor network theory*. Oxford, UK: Oxford University Press.

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

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity*, *43*(2), 265–275.

Little, J. W. (2003). Inside teacher community: Representations of classroom practice. *Teachers College Record*, *105*(6), 913–945.

Little, J. W. (2012). Understanding data use practice among teachers: The contribution of micro-process studies. *American Journal of Education*, *118*(2), 143–166.

Mandinach, E., Honey, M., Light, D., & Brunner, C. (2008). A conceptual framework for data driven decision-making. In E. B. Mandinach & M. Honey (Eds.), *Data-driven school improvement: Linking data and learning* (pp. 13–31). New York, NY: Teachers College Press.

Mehta, J. (2013). The penetration of technocratic logic into the educational field: Rationalizing schooling from the progressives to the present. *Teachers College Record*, *115*(5), 1–36.

Munter, C. (2014). Developing visions of high-quality mathematics instruction. *Journal for Research in Mathematics Education*, *45*(5), 584–635.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics* (Vol. 1). Reston, VA: Author.

Nolen, S. B., Horn, I. S., & Ward, C. J. (2011). Assessment tools as boundary objects in novice teachers' learning. *Cognition & Instruction*, *29*(1), 88–122.

Saville-Troike, M. (2008). *The ethnography of communication: An introduction* (Vol. 14). Indianapolis, IN: John Wiley & Sons.

Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, *26*, 482–496.

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

Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education*, *118*(2), 113–141.

Stringfield, S., Wayman, J. C., & Yakimowski, M. (2005). Scaling up data use in classrooms, schools, and districts. In C. Dede, J. Honan, & L. Peters (Eds.), *Scaling up success: Lessons learned from technology-based educational innovation* (pp. 133–152). San Francisco, CA: Jossey-Bass.

Supovitz, J. (2012). Getting at student understanding: The key to teachers' use of test data. *Teachers College Record*, *114*(11), 1–29.

Van Maanen, J., & Barley, S. R. (1984). Occupational communities: Culture and control in organizations. *Research in Organizational Behavior*, *6*, 287–365.

Wayman, J. C., & Stringfield, S. (2006). Data use for school improvement: School practices and research perspectives. *American Journal of Education*, *112*(4), 463–468.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Erlbaum.

Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, *19*(3), 239–259.

Yin, R. K. (2009). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: SAGE.

Yoon, I. H. (2012). The paradoxical nature of whiteness-at-work in the daily life of schools and teacher communities. *Race Ethnicity and Education*, *15*(5), 587–613.

Manuscript received September 8, 2013
Final revision received January 23, 2015
Accepted January 26, 2015