Educational Policy
http://epx.sagepub.com

Managing in the Middle: School Leaders and the Enactment of Accountability Policy

James P. Spillane, John B. Diamond, Patricia Burch, Tim Hallett, Loyiso Jita and Jennifer Zoltners

Educational Policy 2002; 16; 731
DOI: 10.1177/089590402237311

The online version of this article can be found at:
http://epx.sagepub.com/cgi/content/abstract/16/5/731

Published by:
SAGE Publications
http://www.sagepublications.com

On behalf of:
Politics of Education Association

Additional services and information for *Educational Policy* can be found at:

Email Alerts: http://epx.sagepub.com/cgi/alerts

Subscriptions: http://epx.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations (this article cites 4 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
http://epx.sagepub.com/cgi/content/refs/16/5/731

© 2002 SAGE Publications. All rights reserved. Not for commercial use or unauthorized distribution.

Managing in the Middle: School Leaders and the Enactment of Accountability Policy

JAMES P. SPILLANE, JOHN B. DIAMOND,
PATRICIA BURCH, TIM HALLETT, LOYISO JITA,
and JENNIFER ZOLTNERS

This article investigates how mid-level managers make sense of and mediate district accountability policy. Arguing that teachers' evolving perceptions and understanding of accountability policies are likely to be mediated by school leaders, the authors explore how school managers enact their policy environments, focusing chiefly on the ways in which they construct district accountability policies. Adopting a cognitive or interpretive frame on implementation, the authors illuminate how school leaders' sense-making is situated in their professional biographies, building histories, and roles as intermediaries between the district office and classroom teachers.

DURING THE PAST DECADE, accountability has loomed large in reform initiatives as numerous state and local government agencies have implemented mechanisms that hold schools accountable for student performance (Clotfelter & Ladd, 1996). At the core of these initiatives is an attempt to

---

AUTHORS' NOTE: Work on this article was supported by the Distributed Leadership Project, which is funded by research grants from the National Science Foundation (REC-9873583) and the Spencer Foundation (200000039). Northwestern University's School of Education and Social Policy and Institute for Policy Research also supported work on this article. All inquiries about this research project should be directed to the study's principal investigator, James Spillane, at Northwestern University, 2115 North Campus Drive, Evanston, IL 60208-2615; e-mail: j-spillane@northwestern.edu. All opinions and conclusions expressed in this article are those of the authors and do not necessarily reflect the views of any funding agency or institution.

EDUCATIONAL POLICY, Vol. 16 No. 5, November 2002 731-762
DOI: 10.1177/089590402237311
© 2002 Corwin Press

fundamentally change authority patterns in schools to motivate teachers to more effectively educate America's children. Most research on accountability has focused on the effects of these initiatives on student achievement and, to a lesser extent, their influence on classroom instruction.

Ironically, the role of school principals in implementing accountability-based policies has gone largely ignored. The work of mid-level management is key because accountability levers operate in and through particular schools where they are understood through existing beliefs, experiences, and ways of doing business. Such levers do not exist in a vacuum, and school managers are not passive receptors of their environments. Rather, they enact their environments, that is, they "construct, rearrange, single out, and demolish many objective features of their surroundings" (Weick, 1979, p. 164). In this article, we investigate how mid-level managers make sense of accountability policies. While teachers often encounter district and state accountability mechanisms through media reports, policy directives, and union newsletters (among other sources), their evolving perceptions and understanding of these policies are likely to be mediated through participation in their school community. We use the term *school leaders* rather than *school administrators* because we view school leadership for instruction as distributed among formal and informal leaders in schools (Spillane, Halverson, & Diamond, 2001).

Most policy analysts invoke a relatively clear and popular causal model—rational choice theory—to explain how a state or local government agency can induce an agent (in this case, teachers and schools) to change behavior if they mobilize incentives or sanctions that matter to that agent. However, even if these are the causal mechanisms at play, it is not clear what happens once a school district manages to get a school's attention via the mobilization of rewards and sanctions. Citing rising test scores, Chicago public schools are frequently upheld as a model of successful "accountability" reforms, but the manner in which such policy levers work on and in schools to increase test scores remains something of a black box. Moreover, the available evidence suggests that although student achievement has improved district-wide, these improvements have not been uniform across schools. Whereas some schools have made substantial gains, others have not. Sixty-six elementary schools, more than 15% of the district's elementary schools, had minimal gains or losses in student achievement in the 1999-2000 school year (Hess, 2000). What are the strategies that schools use in response to accountability mechanisms to improve student learning? What does it mean for accountability policies to work at the school level?

Adopting a cognitive or interpretive frame on implementation, we unpack the process through which school leaders implement, and in the process mediate, accountability policies. On one hand, the causal theory in accountability policies is straightforward: If you mobilize tangible rewards that matter to people, they will pay attention and work on the problem you believe they should be working on. However, teaching and learning and leading these endeavors are uncertain. School leaders and teachers operate in uncertain and highly dynamic environments, where means and ends are not always clear. Faced with shifting, limited information, school leaders and teachers do not always act as the profitable or self-maximizing actors portrayed in rational choice models.

Theoretical Frame[1]

Recent scholarship in a number of disciplines investigates the role of implementers' sense-making in the implementation process, underscoring that the reform ideas that implementers construct from policy influence what they do and do not do in implementing that policy (Ball, 1994; Guthrie, 1990; Hill, 2001; Lin, 1998, 2000; Spillane, 2000; Yanow, 1996). Implementation involves interpretation because implementers must figure out what a policy means and whether and how it applies to their school to decide whether and how to ignore, adapt, or adopt policy locally.

A sense-making frame. Human cognition is complex. A key though often ignored aspect of cognition is sense-making, that is, the ways in which people make sense of their environments. Sense-making underscores that "people generate what they interpret;" interpretation is part of sense-making (Weick, 1995). These ideas are not new, forming the basis for the social interactionist tradition in sociology; as Blumer (1969) argued, "Human beings act towards things on the basis of the meaning that things have for them" (p. 2). Weick (1995) argued that sense-making is "grounded in identity construction," and takes place in the context of "ongoing" projects, leading to behaviors that "enact" the environment. Although individuals have loomed large in much of cognitive science, human cognition is not a solo affair. Sense-making is situated, that is, tied to the situation in which stimuli are noticed, interpreted, and subsequently acted on (Brown, Collins, & Duguid, 1989; Resnick, 1991; Suchman, 1988; Weick, 1995). Explanations and actions are enabled and constrained by perceptions of social appropriateness and collective meaning.

A sense-making or cognitive perspective serves as a supplement rather than an alternative to more conventional institutional, rational, and political perspectives on the implementation process in that it expands conventional models by increasing their explanatory power (Spillane & Reimer, in press). Hence, we background the political and institutional dimensions of implementation instead of ignoring them. Politics and the institutional sector are important contexts for mid-level managers' sense-making.

Institutional sector and sense-making. Social agents' actions are situated in institutional sectors that provide norms, rules, and definitions of the environment that constrain and enable action (DiMaggio & Powell, 1991; Scott & Meyer, 1991). These tacit schemata define appropriate structures and give meaning and order to human action in institutional sectors (Scott, 1992). In this scheme, education policy is about preserving the legitimacy of the institution—"logic of confidence"—to maintain public support (Meyer & Rowan, 1977).

Although core original work in the new institutionalism did not present institutional sectors as all determining (DiMaggio & Powell, 1983), subsequent work downplayed the role of human agents, contributing to a more structurally determinist view of human activity. This overemphasis on the role of institutional schemata tends to overlook human agency and other contexts that are potentially important for implementing agents' sense-making.

Institutional schemata do not determine sense-making as much as they obliterate human agency. Our cognitive perspective acknowledges that although institutional arrangements structure human sense-making they do not obliterate human agency, which shapes these arrangements (Giddens, 1984). Hence, we attend carefully to the influence of local agents in constructing policy messages.

Political circumstances and sense-making. Political arrangements are also an important context for mid-level managers' sense-making; they face major political challenges because their position in the organizational hierarchy focuses their work in at least two directions. On one hand, school leaders are street-level workers dependent on and responsible to their local community stakeholders and the district office for implementing school policy (Lipsky, 1980). On the other hand, school leaders depend on other street-level workers—classroom teachers—for the successful implementation of these policies.

Relations between managers (both district office and school) and street-level workers (both classroom teachers as well as school managers vis-à-vis district managers) are characterized by dependency and conflict (Lipsky, 1980). Actors at each level of the school system operate in relatively independent political arenas or "games," mobilizing resources that can be used to advance, sabotage, or simply ignore the efforts of actors at other levels (Firestone, 1989). The interests and objectives of mid-level managers and those of street-level workers are often in conflict with street-level workers interested in minimizing the dangers and discomfort of the job and maximizing income, personal gratification, and their autonomy. Managers seek to achieve results consistent with agency objectives and work to restrict street-level workers' autonomy and discretion. However, school principals' formal authority with respect to teachers is limited, with principals relying on more subtle strategies including the manipulation of language and ideological control (Anderson, 1991; Dreeben, 1970; Lortie, 1975).

The autonomy that teachers and other street-level workers exercise is in part a product of teachers' isolation and conditions which leave them to practice alone with little direct supervision (Lortie, 1975). Street-level workers have some important resources with which they can resist managers' directives. Teachers' expertise and their willingness to get involved in the school's agenda are critical resources. Teachers use a variety of strategies to influence administrators including bargaining, bluff, threats, flattery, and exchange (Bridges, 1970; Fraatz, 1987). Hence, classroom teachers enjoy relatively high degrees of discretion and relative autonomy from organizational authority, contributing to school leaders' dependency on teachers.

Yet at the same time, teachers are partially dependent on principals and potentially other school leaders who allocate funding, curricular materials, and class assignments. School leaders' sense-making is situated in this set of political relations in which different actors and groups of actors use power to influence each other and preserve their own interests and agendas.

Our sense-making perspective is different from the political perspective in that we do not assume that local actors are motivated chiefly or exclusively by self-advancement. Indeed, some research in political science suggests that bureaucrats tend to be hard working and that they do not typically work to undermine policy or directives from above to advance their own self-interest (Brehm & Gates, 1997). Furthermore, the political perspective emphasizes how actors select alternatives that advance their own interests. To make such selections, local actors must first figure out what a policy or innovation means to decide whether and how to ignore, sabotage, adapt, or adopt it to suit their self-interests. Political arrangements are an important context for local actors' efforts to make sense of policy.

RESEARCH METHODOLOGY

This article is based on data from the Distributed Leadership Project, a 4-year longitudinal study of elementary school leadership funded by the National Science Foundation and the Spencer Foundation. The first full year of data collection (Phase 1) began in September 1999 and focused on eight Chicago elementary schools. Year 1 data collection involved between 50 and 70 days of fieldwork in each of our eight sites. The three schools that are the focus of this article were selected because they represent important variation among schools in our study in response to accountability policies. We used a theoretical sampling strategy (Glaser, 1978; Glaser & Strauss, 1967) to select schools based on poverty level, racial and ethnic composition of student population, degree of improvement in student achievement, and variation on measures of school improvement including "academic press," "professional community," and "instructional leadership"[2] (process measures) and "academic productivity."[3]

Data collection. Research methodologies include observations, structured and semi-structured interviews, and videotapes of leadership practice. Researchers spent the equivalent of 3 to 4 days per week per school during a 10-week period in the fall of 1999 and a 12-week period in the spring of 2000. Leadership events observed included grade-level meetings, faculty meetings, school improvement planning meetings, professional development workshops, supervisions of teaching practice, homeroom conversations, and lunchroom conversations.

We completed interviews with teachers at the second and fifth grade levels and school leaders (including lead teachers). Interview protocols focused on school leaders' agenda and goals, their responsibilities, and the key tasks they perform as part of promoting instructional change in mathematics, science, and literacy.[4] We also conducted postobservation interviews with school leaders about specific instances of their practice that we observed.[5]

Data analysis. Data collection and data analysis were closely integrated, allowing us to examine patterns and working hypotheses as they emerged from data analysis and refine data collection strategies (Miles & Huberman, 1984). A commercial computer-based qualitative coding program—NUDIST—was used to code all project data. NUDIST allowed us to code the emerging ideas and concepts from the data into free nodes that can be compared and related to each other, forming larger "parent" nodes that can be stored in an index system that brings the different components of the project together. Coders worked together to code transcripts initially to develop a shared understanding of what each code meant. Once coders had developed a "taken as shared" understanding of these codes, they worked independently.

ACCOUNTABILITY POLICY: THE CASE OF CHICAGO

Accountability policy initiatives involve at least two components: specific student performance outcomes and rewards and sanctions for schools (Clotfelter & Ladd, 1996; Elmore, Abelmann, & Fuhrman, 1996). Student performance outcomes as measured by tests rather than inputs (e.g., number of certified staff) are the primary mechanism that states and school districts use to hold schools accountable. Although the 1988 Chicago School Reform Act (P.A. 85-1418) included the decentralization of decision making to the school site level and the formation of Local School Councils (LSC), the Chicago School Reform Amendatory Act of 1995 gave much authority to the chief executive officer (CEO), appointed by the mayor, who was able to place poorly performing schools in remediation or on probation based on their performance on the Iowa Test of Basic Skills (ITBS). Student performance on the ITBS at benchmark grades became the district's primary measures of school accountability and progress.

The second component of most accountability measures involves the creation of a system of rewards and sanctions as well as intervention strategies designed to motivate schools to improve student achievement. In Chicago public schools, the key sanction is the power of the CEO to place schools on probation because of low performance. For example, in 1996 the CEO placed 20% of the elementary schools, 109 schools, on probation because fewer than 15% of their students performed at or above national norms on the reading and mathematics sections of the ITBS (Hess, 2000; Wong & Anagnostopoulos, 1998). Schools on probation are required to develop a supplemental school improvement plan that outlines specific strategies the school will take to improve student achievement and defines criteria that will be used to judge the school's progress toward improvement. For technical assistance, schools use their discretionary funds to purchase the services of an external partner whom they can select from a district-approved list. If the district decides that a school has not made adequate progress, the CEO can have the school reconstituted, ordering new LSC elections and replacing the principal and faculty.

Some efforts to transform accountability arrangements also include rewards or sanctions for students. In 1996 the school district also ended social promotion, informing the students that beginning with the 1996-1997 school year if they failed to achieve at a certain level on the ITBS they would have to attend summer school, and if by the end of the summer they still failed to achieve at the required level they would not be promoted.

MANAGING IN THE MIDDLE: SCHOOL LEADERS AND DISTRICT ACCOUNTABILITY POLICY

Inattention and perfunctory attention are pervasive problems when it comes to the implementation of public policy (Firestone, 1989). The problems of inattention are likely to decrease, however, if the policy-making agency manages to mobilize tangible incentives that matter to implementing agents and agencies. Hence, when faced with the threat of closure or probation due to unacceptable student test performance, as defined by the school district, schools are likely to pay attention to policy proposals. Still, attention to policy is a complex matter (Spillane, 1999). Although schools may attend to district accountability measures writ large, the ways in which they come to understand and hence what they attend to in an effort to implement these policies may vary substantially across schools. To explore these matters we use three mini-cases that focus on how school leaders at three Chicago elementary schools made sense of district accountability efforts.

Baxter Elementary

At Baxter Elementary, administrators and teacher leaders have come to view standardized test data as an integral tool for helping teachers improve instruction in particular content areas. The principal has led the effort to run test score analyses on data provided by the district so that teachers can chart the progress of school-level reforms by subject area and grade level. These leadership strategies have shifted the import of test score data within the school's improvement agenda—transforming student outcome data from something the district simply demands to a tool that the school fully expects to use.

Sociohistorical Context

The focus on test scores is somewhat ironic. Baxter traditionally has little reason to worry about student test scores and district accountability levers in general. The school has met national norms in core subject areas and at the benchmark grade levels (third, fifth, and eighth) for at least 5 years, has received national recognition for academic excellence and leadership, and is locally recognized as one of the best in the city. In 1998, 60.4% of students scored at or above national norms on the ITBS in reading, whereas 69% did so in mathematics. Baxter's achievements come as the neighborhood that surrounds the school undergoes dramatic demographic changes. Once a White Jewish enclave, the neighborhood is now occupied by families from Eastern Europe, the Middle East, and parts of Asia. Meanwhile, the White and Black student population has decreased by 7% and 2%, respectively, the Latino population has increased by 7%, and the Asian population has increased by 2% during the past decade. The percentage of low-income students increased from 24% to 66% between 1990 and 1998.

Sense-Making and Instructional Improvement

However, although the school faces no real threat from the district's accountability policy, test score data represent an important instructional tool under the school's current leadership. In the context of staff development and curricular meetings, principal and teacher leaders send teachers a consistent message about the role of test scores within the school's own instructional agenda: For our own sake, let's use it! In particular, by using disaggregated test score data as an ongoing reality check on school progress, Baxter's school leaders have responded to district accountability measures in ways that enable them to build support for ongoing curricular and instructional improvement efforts. Discussion of test score data has become a regular feature of management and grade-level meetings, helping guide discussion of curricular priorities. For example, as part of its decentralized governance structure, the school holds periodic Saturday meetings. The agenda for one such meeting was to create a schedule for developing the annual school improvement plan required by the district. The principal and other teacher leaders pressed to hold off on meetings until they had standardized test data back from the district, even though beginning the process earlier would have been more convenient.

> After Mr. P finishes his argument for spreading out the School Improvement Planning [SIP] process over the course of the year, S asks, "Why can't we get the SIP done early rather than waiting until spring when we will be busy with testing . . . that way we'll have more time and won't be so pressured." Mr. P is quick to respond, and there is head nodding from several other teachers in the room. "Well, because we use the data from the standardized tests to develop the SIP and that's not available until February. "But why can't we use old data from the SIP?" "So much has changed," D, a veteran teacher, responds, "we need to look carefully at the new stuff otherwise we'll be doing things we don't really need to or ignoring things we should." (field notes, meeting)

The field notes exemplify the ways in which Baxter leaders work to shift teachers' understanding about the relationship between district accountability pressures and existing school improvement strategies. The principal's proposal is not simply an issue of scheduling; it is an argument for leveraging district pressures to work for rather than against school instructional plans.

Baxter's principal explains that the need for better and disaggregated data on student learning became more pressing as faculty leadership for curricular change gained momentum and faculty began to get much more assertive and active in their curriculum development needs. In response, the principal, assistant principal, and new dean of students began to run trend analyses of standardized test data to give teacher teams a clearer picture of the institutional problems that the faculty teams were trying to address. Faculty assumed that Baxter was doing a much better job in educating its students than other schools. Trend test score analysis was intended to challenge this assumption and get faculty to recognize the problem. The principal explained it like this:

> "Ok, I see we have a problem, what kinds of things should this school be doing to address the problem," explains the principal. "The analysis made clear that out of the 12 schools, Baxter was either at the bottom or really close to the bottom, in terms of the amount of actual growth that students were making. Forget about where the growth started, forget about the base. Forget about the end. Just, you know, how many months of progress, on an average, were sixth graders achieving at Baxter School between 1991-1995." (interview, principal)

At Baxter, school leaders' efforts to maintain credibility and distinction as a star school are a powerful motivator for attending to district accountability levers. Although the principal is fondly known in the school for ignoring the more mundane of district requests for school-level information, he and other leaders are also described as highly active in employing data on Baxter's improvements on test scores to win funds and build local support for his initiatives. For example, school leaders have compiled their comprehensive longitudinal breakdown of school test score improvements (described above) and included it in proposals and presentations with potential granters, research organizations, and local governance councils.

School leaders also take pains to repackage standardized test data in ways that will capture teachers' attention. Standardized test scores from the district typically arrive via cardboard box and massive spreadsheet. At a recent faculty meeting, this same information was reorganized into relatively easy-to-read charts and color-coded by grade level.

> When I entered the auditorium, S motioned me to several neat piles of materials on the auditorium stage. These were the materials for the meeting. Most of the faculty meetings I have attended have a cryptic agenda at best. This meeting was different. Not only did the materials display detailed information on student outcomes, but the charts [that I learned later Mr. P had developed—and that were a mainstay of meetings he organized] were easy to read with labels for those of us who had difficulty with

> numbers and charts. Different kinds of data were displayed on different kinds of charts and each chart has its own color. The first chart was already up on the screen—projected by the overhead that I assumed Mr. P would use for his presentation. (field notes)

This scenario illustrates the somewhat overlooked role that principals play in relationship to teachers' understanding of district accountability pressures. Here the homegrown charts aimed at not just informing teachers about district requirements but helping them see the relationship between external demands and more internal instructional concerns.

Mr. P. and other school leaders use district-standardized test data as a lever not only to motivate faculty to work for improvements on standardized tests but to build momentum into the school's distinct and collective vision for reform. Baxter's approach to school improvement planning offers a window on how this works. A typical Chicago approach to school improvement planning, explains a seventh and eighth grade teacher who has spent several decades working as a teacher in the system, is to simply take last year's school improvement plan and tweak it, making minor alterations in focus and substance (interview, teacher). In contrast, at Baxter, standardized test data released in the spring are used as the focal point for grade-level and schoolwide discussions around the plan. For example, spearheaded by the principal, Baxter's leadership committee undertook careful study of school-standardized test data in math. They wanted to know how and whether high math scores in Grade 3 were being sustained through Grade 5. Through longitudinal analysis, the team determined that somewhere in Grade 6, students' scores started to slump. Suspending school convention, the leadership committee convened a joint task force of two groups that traditionally had little interaction: the third and fourth grade teachers and the fifth and sixth grade teachers. This joint committee met for the good part of a year and as a result of their work tried to build greater alignment in math topic coverage across grade levels.

Nature of Leadership

The school's approach to standardized test data and the in-school processes that they use in this approach seem very much rooted in the individual and collective identities of school staff. Mr. P, the principal, is described by other faculty and himself as "very much of a data guy." A former elementary school teacher, Mr. P is a man who peppers conversations with outside visitors to his schools with reference to statistical studies of education reform and whose desk overflows with the trend analyses of student performance he generates for faculty.

Mr. P's passion and skill at number crunching made him a good candidate for the principalship of a school that was under pressure to improve its compliance with district reporting requirements. However, he was drawn to the school for a different set of reasons. Fifteen years ago, Mr. P was a principal at a suburban school in a neighboring state, having started his career as an elementary school teacher. He resigned what he describes as this "plum principalship," and moved himself and his family to Chicago in 1989 largely because of the Chicago decentralization reform which was then in its infancy. The Chicago Reform dismantled the central office and devolved decision-making authority over budget and curriculum to the school level. Mr. P explained,

> Professionally speaking, what interested me here was just this very exciting concept of being able to function as a principal working for a local board, having total accountability to that local board and also being able to work with them to set our own personnel and program policy.

When Mr. P arrived at Baxter, he, in collaboration with other leaders, organized teachers into different work groups and assigned them the responsibility for developing or improving curriculum in different content areas.

We rarely think of school-level leaders as "data crunchers," but at a school like Baxter, where meetings typically begin with the question, "What information do we have on this problem?" Mr. P now finds himself in good company. In 1999-2000, staff members (including the assistant principal, the dean of students, literacy specialist, and grade-level teacher leaders) volunteered to develop the school's annual district-mandated improvement plan. The group decided early in the process that the data collected by external state monitors the previous year was insufficient. To address this problem, they designed their own survey and classroom teacher interview protocol, a protocol designed to collect comparable data across teachers on classroom practice and staff development needs. After months of work, this data was compiled, carefully analyzed for cross-grade and grade-level patterns, and then compiled into a 20-page report complete with bar graphs. Indeed, as an organizational community, the school could rival a modern day corporation in its complex layered decision-making structure and highly regulated systems for making sure every teacher has input and that the input is gathered and aggregated. Every grade is part of a cycle; every cycle has an operations and a curriculum chair, and these chairs in turn make up the membership of the school's leadership committee. It was the leadership committee that was responsible, among other things, for reviewing the report compiled by the school improvement planning team. As part of this process, the grade-level leaders who composed the team reviewed and discussed the report's findings at one of the school's weekly grade-level meetings to cross-check for accuracy and gather supplemental data.

Conclusions

The product of these and other efforts, by one teacher's report, is a curriculum that "is getting closer and closer to a standards-based approach and teaching students what they need to know." In addition, in a school district that is moving toward standards-based indicators in its standardized tests, this may increase the likelihood of maintaining good test scores. However, whereas Baxter's administrators and teacher leaders tend to view standardized test data as a helpful tool in the school's improvement efforts, Baxter's teachers do not necessarily share these views. The majority view the tests as an annual inconvenience, with some teachers reporting that they try simply to ignore them and structure course content as though the tests do not exist. Other Baxter teachers display more strong and negative reactions. In the spring of 2000, when the district began requiring second grade teachers to pilot a new standardized test, a group of Baxter's second grade teachers wrote a searing letter to central office management. The letter argued, among other things, that administration of the test (which in this instance was individualized) prevented them from covering topics the second graders would need to succeed in third grade.

Teachers' skeptical views of the tests make it critical for Baxter's leadership to frame the district's accountability pressures in the right light—not as a recipe or blueprint for action but as a starting point for school staff to reflect on their practices and develop alternative strategies. Indeed, Mr. P, in describing his ingredients for school success, is quick to critique those aspects of the district's accountability policies that he feels undercut teacher autonomy. In large and small group meetings, he and other Baxter leaders frequently identify teacher leadership and input into the curriculum as the primary source of and influence on Baxter's improved test scores. The principal explained, "The more chances people have to talk, and not just talk in circles but talk about common ground problems, the more comfortable they feel with change, and the better they get in their practice" (interview). They do this not only with outside audiences but also in the context of internal faculty meetings and school-wide memorandums (field notes, faculty meeting). In these and other ways, they try to answer to the district and use district pressures to pursue leadership goals without compromising the principles of their vision and their relationships with faculty. So Mr. P, although openly critical of many aspects of district policy with his teachers, reframes accountability policy for his staff and manages to use district test score data as the basis for a dialogue about improving instruction.

Dodge Elementary

Dodge is a poignant example of how complicated sense-making contexts problematize straightforward efforts to enact school reform. As she formulates her instructional leadership efforts, the principal (Ms. Wen) struggles to reconcile the opinions and perceptions of different groups with her own views on accountability and Dodge's standing relative to other schools.

Sociohistorical Context

Located within a 10-minute drive of Baxter, Dodge Elementary likewise serves a diverse student population: 39.2% White, 30.9% Asian/Pacific Islander, 21.6% Hispanic, 8.2% African American, and 0.2% Native American. Multiple immigrant groups have settled in the neighborhood, yet as these new families gain prosperity, they often move out of the area. The student population at Dodge is constantly changing, increasing from around 1,000 students in 1990 to approximately 1,600 in 1999. The school administration has also undergone a period of flux, with 4 principals in the past 9 years. Despite the upheaval, student achievement on the ITBS is high compared to most Chicago public schools, and ITBS scores increased gradually from 1991 to 1999. In 1999, 53.3% of the students scored at or above the national average in reading, and 59.3% of the students scored at or above the national average in math. Dodge school is in no danger of being put on probation in accordance with recent accountability-based policies.

Instructional reform at Dodge reflects efforts of the different administrations to make sense of accountability, the first change coming with the retirement of the veteran principal, Mr. Welch, in 1992. The staff loved Welch, in part because of his resistance to external policy levers threatening the established practices that the staff perceived as successful. In particular, Welch resisted tinkering by the newly formed LSC. The LSC chair reflected, "I think curriculum issues were never the purview, were never allowed to be the purview of the LSC. In fact, I was actively discouraged" (interview). When Welch retired, the long-time assistant principal took his place for 1 year before choosing early retirement.

Welch's retirement opened the door for the LSC to hire Ms. Jackson, an outsider to the Chicago public school system with a disposition for change. A former principal in a nearby suburb, Jackson emphasized direct instruction in reading and the University of Chicago's "Everyday Math" curriculum. Jackson also reorganized the fifth to eighth grades from self-contained classrooms into departments, and hired an outside consulting firm to align the curriculum with the newly formed Chicago public school standards. However, many of the veteran staff members resented Jackson's lack of experience in the Chicago public school system and her efforts to change methods they felt were proven, and continued on with their own ways of doing business. They paid little attention to the new curriculum, rarely mentioning it or emphasizing the fact that it sits dormant in their drawers. Just prior to the start of the 1997-1998 school year Jackson abruptly quit. Dodge began the year without a principal, giving the staff full autonomy to (re)establish their own practices.

In the middle of the school year (start of 1998), the LSC hired Ms. Wen. Originally from Hong Kong, Ms. Wen came to America in the eighth grade and became a teacher when a guidance counselor encouraged Wen to obtain her teaching certificate. After teaching regular and bilingual classes at Catholic and public schools while obtaining an English as a second language (ESL) endorsement and her type 75, Wen sought a "new challenge" and moved on to assistant principal positions at two small elementary schools. Then she was encouraged by one of her college professors to apply to a new Principal Leadership Training and Education (PLTE) program. Ms. Wen was accepted into the program and then selected to act as an associate principal at a high school that had been placed on academic probation by the Chicago public school system. In time, Wen's mentor in the PLTE program encouraged her to apply for the position at Dodge. Although Wen did not expect to be hired, the LSC chair was impressed with her PLTE credentials, bilingual experience, and energy. Ms. Wen was hired over the opposition of the LSC teacher representatives and the staff they represented.

Ms. Wen started work on December 21, 1998. In an attempt to find her bearings, Wen called on the long-time assistant principal, Mrs. Hecky, for help. However, it was winter break and Mrs. Hecky, on vacation and in recuperation from breast cancer, did not come in. Their relationship soured, and Ms. Wen eventually replaced Hecky with an old colleague. The veteran teachers, suspicious of Wen's blunt style, lack of experience as a principal, and inexperience at large schools, chafed at Wen's treatment of Mrs. Hecky, and braced themselves to defend the individualized practices they had established in the days of Mr. Welch.

Sense-Making and Instructional Improvement

The history of Dodge and the political atmosphere play an important part of the sense-making context as Wen struggles to enact Chicago public school accountability policies. The sense-making context has three main elements:

(a) how the LSC and staff interpret shifts in neighborhood demographics in relation to test scores and accountability; (b) Wen's struggle for legitimacy as a new principal and her own interpretation of test scores and accountability policy, especially in comparison to nearby schools such as Baxter; and (c) the history of teacher autonomy at Dodge.

An important part of the sense-making context at Dodge involves how school staff and the LSC construct demographic shifts in relation to accountability. During the past decade the Caucasian student population dropped from 51% to 40%, whereas the African American, Latino, and Asian populations grew to 2%, 6%, and 3% respectively. However, the most dramatic shift has been in the percentage of low-income students, which changed from 44% in 1990 to 73% in 1998. Thus, when the scores on the ITBS slipped in reading and math, the LSC was content with an interpretation that the drop was statistically insignificant and could be explained by changes in student demographics. The LSC does not correlate lower scores with the administration they selected. Many of the veteran teachers at Dodge also interpret declining scores in light of changes in the student body. However, where the LSC uses this interpretation to validate their principal hire, some (although not all) veteran teachers use this interpretation to validate established instructional practices. At a staff meeting, one teacher stated, "We're getting more and more kids now with problems at home. . . . There's no discipline in the household, and I can model things here, but if they don't get it at home" (field notes). Teachers like this understand declining test scores as a function of shifting demographics beyond their control, and believe that alternative teaching strategies are unlikely to remedy the situation.

A related part of the sense-making context involves Ms. Wen and her struggle for legitimacy in the eyes of the staff combined with her own interpretation of the scores. Wen has spent considerable time cultivating legitimacy with a veteran staff, many of whom view current scores as reasonable and a validation of their practices. Yet, unlike the LSC and some of her staff, Wen does not believe changing demographics are an acceptable explanation for declines, especially because student high school choice has become linked to scores. She explained,

> But when I look at the test results, fifty percent are succeeding, I look at it the other way, fifty percent of our children are not succeeding. . . . Bottom line is the kids have to bring those grades up to apply for the best high schools. They're not going to be considered if they don't have the scores in their hands. (interview)

For Ms. Wen, test scores matter, especially in relation to high school opportunities.

Wen's concern about scores is often expressed in light of other area schools, especially Baxter. Take this interaction between Wen and some second grade teachers at a grade-level meeting:

> Ms. Wen tells them the school needs to do something to improve reading, because their scores are down "1.3" on the Iowa tests. In contrast, the reading scores at Baxter are at 70, "I have to go over there." A teacher responds, "I'll go with you," adding, "They must be teaching to the test" because the two schools are "servicing the same population." (field notes)

Although Wen lends weight to the test scores where the LSC and some staff do not, she does so in regard to the scores at another area school, which in turn affects how she presents the material to the second grade teachers. Given that her success will be gauged to some extent based on comparisons to schools with similar populations, Wen's attention to Baxter is not surprising.

It is also important to consider Ms. Wen's difficult transition into the role of principal and her attempts to create legitimacy with her staff. Whereas the LSC has the power to hire and renew her contract, the teachers have the ability to make Ms. Wen's daily life miserable, and many who are unhappy with her appointment have made such attempts. Ms. Wen referred more than once to the difficulties she has encountered in her role as principal: "I think that this [principalship] is probably the most difficult task that I have ever taken on. So I just try to survive every day. I just look at every day as a new day" (interview). At one point, a group of teachers persuaded the district to appoint an external evaluator based on their formal complaints about Ms. Wen. These circumstances, in part, reflect the laissez-faire style of previous administrations and efforts by Wen to provide more structured and centralized leadership. Ms. Wen explained,

> One teacher said to me one day in the corner, and I think it really gives you a lot of information, she told me that . . . they'd [teachers] been running the school without a principal for 6 months. . . . Well, that's not going to happen with this administration. . . . So I think some of the resentment come from the change of requiring teachers to take a little bit more of a step of complying—doing things. (interview)

In addition to such conflicts, Dodge lacks Baxter's collaborative approach to instruction. At Dodge, teachers work largely on their own to craft instructional approaches and strategies enabling students to do well on the ITBS. Consider the following remarks from teachers:

> We tried to also work out of the IGAP or ISAT books. The test books. They are just preparation for the ISAT and Iowa's. So a lot of it was making up my own kinds of work for the students. (interview)
>
> As much as I hate to say it I am teaching towards the test. Not something I like, but come the end of the year and the students need to get a 7.8 in order to graduate. . . . Until that changes, I don't see anyone not teaching towards the test. How you do that is the big question. How I do that, I like to get the students involved in groups and I think that grouping them up with one student that might be the leader helps out a lot. (interview)

These comments underscore the individualized nature of the work at Dodge, where joint efforts to devise instructional strategies were rare compared with Baxter, despite Wen's efforts.

Nature of Leadership

These complex features of the sense-making context make efforts by Ms. Wen to respond to district accountability mechanisms all the more challenging. Concerned about students' performance, Ms. Wen believes that the key to improvement has to do with transforming teaching practice rather than in introducing new curricula. To quote Ms. Wen, "I think a lot of time is wasted in deciding what curriculum should be taught and more emphasis should be put on refining the way of teaching" (interview). But Ms. Wen's desire to use accountability levers to improve test scores are complicated by how the LSC and the staff interpret shifts in student demographics and test scores, as well as attempts by many teachers to undermine her leadership.

Even so, Ms. Wen is trying to press change in instruction without alienating teachers, and she believes scores would improve if teachers would increase focus and collaboration: "A lot of time is wasted in not focused instructions in this school. But there is some good teaching going on. But the problem is there's not much of collaboration so teachers tend to deal with the closed door and do what they do" (interview). The problem is complicated further by the fact that many teachers at Dodge believe that probation policies are intended for schools performing poorly, not for Dodge where, despite a decline in test scores, more than 50% of students score above national norms. Under these circumstances, many teachers at Dodge find it difficult to see the need for change. Consider the following example from a staff meeting:

> Then Ms. Wen turned the floor to Mrs. D [upper grade literacy teacher], who had gone to the staff development on the structured curriculum. Mrs. D began, "First of all, people were really angry at the meeting because they spend so much materials on this, but basically" it's just lesson plans aligned with state goals. Then a teacher asked, "I thought this was just for schools on probation." And another teacher who had been to

> the staff development answered, "It's not mandated except for schools that are on probation." Then the first teacher repeated her feelings "for those of you who have been in Chicago public school before," "it's just like the old punch cards" and a "waste." But Mr. W disagrees, saying, "I don't think it's a waste," that during summer school it laid everything out and you can still "work creativity into place." Mrs. D agreed, saying, "Yes, it is a good resource" (not a total waste), and Ms. Wen said, "It's not mandated, we are told it is for the 1st year teachers [next to me a 1st year teacher makes a "hrmph" sound]. So how do you use it? It's up to you." (field notes)

At Dodge many teachers' understanding of the district accountability effort is influenced by their perception that these policies are for schools on probation, not for a school like Dodge where 50% of students perform at or above national norms. Cognizant of this, Ms. Wen is reluctant to use district accountability policy and test scores to leverage change at the school. For example, in the excerpt above, Ms. Wen, towards the end of the meeting, agrees with teachers that the district material is indeed optional, and it is up to other teachers whether and how they use it. Cognizant of teacher resistance coupled with her dependency on them for her legitimacy as a leader, Ms. Wen has to manage a delicate balance between district policies and her desire to improve test scores, on one hand, and the opinions of classroom teachers on the other. Given the complexity of the sense-making context, it is not surprising that little instructional change has occurred at Dodge, despite the widespread rhetoric of accountability.

Conclusion

In sum, the sense that Ms. Wen makes of district accountability is a function of multiple overlapping contexts including Chicago School Board policies, the demographics of the neighborhood, the LSC, Wen's own beliefs and struggle for legitimacy, and the teaching staff's practices and beliefs. Each of these elements makes up a portion of the "sense-making context" at Dodge school. Nestled inside this context, Ms. Wen struggles to make sense of policy initiatives while formulating the best route for instructional leadership, a challenge to say the least.

Waxton Elementary

Sociohistorical Context

District accountability policies figure prominently at Waxton Elementary School. That is not surprising, given that the school was on academic probation until fall of 2000. In 1993, fewer than 8% of Waxton students scored at or above the national norms on the ITBS reading test and only 11% did so in mathematics. In 1999, the situation had improved some, with 17.6% of the students scoring at or above the national norms in reading and 23.7% doing so in mathematics. Unlike Dodge and Baxter schools, the student population demographics at Waxton have remained stable. Waxton's students are 100% African American and in 1998 97% of the students were from low-income families based on eligibility for free or reduced-cost lunches. These numbers have not changed significantly in the past 10 years. Although the population demographics have remained largely unchanged, the school has a high annual student mobility rate of just greater than 40%. Therefore, the students who make up the school change substantially from year to year.

The policy pressures at Waxton are somewhat different from those found at Dodge and Baxter. At those schools, students have performed well on the ITBS and little external pressure is exerted to press for improved student outcomes. Waxton, on the other hand, faces clear pressures to improve. In fact, because of its probation status, the office of accountability appointed the school's new principal, Dr. Sampson, just prior to the 1999 school year. As she explained,

> One day I received a call from the Office of Accountability asking if I was interested in an appointment. I went downtown to Accountability for an interview and the week before school was to formally open I was asked to come to Waxton.

Understanding how the school responds to external accountability is best accomplished by paying close attention to several features of the sense-making context: the school's probation status, external pressures for improvement in student outcomes, the appointment of a new principal, and the principal's efforts to establish legitimacy in her new role.

Sense-Making and Instructional Improvement

Although student test scores at Waxton have increased during the past 7 years, this increase has not been large enough until fall 2000 to get Waxton off of the district's probation list. The school's position in relation to the district's accountability regime has shaped its response to high stakes testing. The prominence of the testing and probation issues in the minds of school leaders became apparent early in our research involvement with the school. When asked about her goals for the year, the principal explained that "the obvious goal is to get off probation! Now that's it in a nutshell." This external pressure to change students' outcomes was reflected in interviews with other school leaders as well. As the school's language arts coordinator explained,

> The system is assessment driven. . . . The whole Chicago school system. That's all you hear about on the media. You know, the test scores. The test scores. So naturally the

> teachers are . . . focusing on, you know, on making sure their children do as well as possible.

The district push for accountability, and the school-specific pressures related to probation status, were important components of the sense-making context and influenced school responses to accountability policy. Three clear examples of this were the leadership focus on benchmark grades in professional development activities, its emphasis on reading as opposed to other subject areas, and the alignment of its curriculum to state and district standards.

Testing and probation loom large in the administration's menu of professional development for teachers. Some 50% of the eight professional development meetings we observed during the year were largely or entirely focused on some aspect of testing. These discussions included topics such as skills tested in language arts, skills tested in mathematics, constructing multiple test items, and preparing students for the ITBS. An interesting twist on these efforts, however, was that they were targeted on certain benchmark grades connected with student promotion. Waxton's administration provided the grade levels that were tested with more attention. For example, whereas Waxton's probation partner provided a few general professional development sessions that targeted the whole faculty, their one-on-one mentoring efforts were focused entirely on teachers at Grades 3, 5, and 8—testing grades. The language arts coordinator explained this focus prior to one visit:

> Teachers from those Grades 3, 5, and 8 [from our external partner school] are coming here today and they're going to observe in those rooms to see how they can assist the teachers in preparing for the test.

Furthermore, the teachers of Grades 3, 5, and 8 were given a set of test preparation books to use to prepare their students, whereas teachers of other grades were not. Therefore, testing influenced the focus of the school's professional development activities by making test preparation a prominent activity and targeting the school's activities at certain grade levels.

In addition to this focus on certain grade levels for teachers' professional development and test preparation support, school leaders at Waxton were also targeting their instructional change efforts at a particular subject area—language arts. As the principal explained,

> Being very honest, language arts, specifically reading, is one area that could impact probation and since the school had been on probation for so long we felt a need to address that curriculum area. And the mathematics scores were slightly higher than the reading so that gave us the second reason.

State standards and district frameworks also figured prominently at Waxton. According to the principal, the standards and frameworks played a key role in focusing the school's curriculum:

> In all three areas we address the standards and the framework statements. [In] language arts we focused of course on all four areas but reading was number one, writing number two. We did emphasize speaking and listening but not quite as much. We plan to make more efforts next year to incorporate speaking and listening. In mathematics it took a second seat to language arts. We realize that problem solving was a much more difficult area than computations mainly because problem solving involves reading. So teachers were made increasingly aware of the need to critically read or to critically think as one was solving the problems.

Waxton long-term probation status has contributed to putting test and test scores front and central in the school's improvement efforts. Although there was a general focus on instructional changes and aligning curriculum to state and district standards, the immediate external pressures of accountability led to a focus of instructional change efforts on certain grade levels and subject matter areas rather than to the school as a whole.

Nature of Leadership

Waxton's approach to accountability policy is shaped by the identities of school leaders, particularly the principal. Dr. Sampson, an African American woman in her 50s, comes from a long line of teachers in her family and discusses her career choice as if it were a given:

> I almost think I was born into it. I had aunts who were teachers, my older sister was a teacher, my older brother was a teacher. . . . It was encouraged. Both my parents encouraged school teaching. . . . I didn't consider anything else. Beyond 8 years old and up . . . I've never considered anything else.

Her educational experiences emphasized teaching with a strong emphasis on curricular issues. She received an undergraduate degree in elementary education and a master's degree in science education. After working as a math-science coordinator at another local school, she went back to school to "get her credentials" as a principal. She then became an assistant principal at her former school, working closely with a principal identified by many as one of the most outstanding in the district. She also participated in a principal training program at a local university. She credits this experience with showing her the importance of balancing the focus on curricular change with an emphasis on the politics of leadership. This combination of curricular knowledge and awareness of the political dimension of school change efforts informs how she makes sense of accountability pressures and seeks to shape instructional reform.

Although the principal is clear regarding her overall goal "getting off probation," one of the key foci of her instructional leadership efforts was to build morale. This was done in several ways. One strategy that the principal used was to send thank you cards to teachers at the beginning of the school year. She explained that "the first month I was here I found a reason to send everybody a thank you card for something." As one teacher said of this practice, "It makes you [teachers] kind of work harder." Dr. Sampson believes that this helped "to boost morale," which was "really one of the major things this year."

A larger scale part of this morale-boosting approach—the pep rally strategy—focused on teacher motivation and was tied directly to accountability. Consider the first meeting between Waxton's staff and the school's new probation manager:

> There was clapping of hands as Beatrice stood up to speak. She began by saying "it is possible to get off probation" (and there was another loud applause) and "you are going to get off probation" (another loud applause). Beatrice then gave a little talk about how she thinks the staff and children at Waxton are capable and can achieve. "Sometimes," she continued, "it's only a question of knowing what to focus on, getting children to be ready for the tests" and getting the right tools.

This approach was evident again at another meeting with the probation manager:

> Beatrice quickly read off the next activities on her list (noting that they were running out of time): "Follow up in the classrooms after discussing each strategy," "test taking—we will do breakdown of test skills, so you focus on the right skills that are asked in the tests," "We did this at N School and found it very helpful" ("yesses" from teachers and interjections of "that's what we need!"). Beatrice's colleague emphasized that Waxton will "get off probation" (another applause).

Whereas these leadership efforts focused on morale building, others were designed to enhance teachers' capacity to transform their classroom practices. The first step of this capacity-building process was the suspension of teacher evaluations for the first quarter of the school year. Dr. Sampson argued that she wanted to first focus on developing teachers' capacity before doing any teacher evaluations. She explained:

> I received so much negative press from everyone about what this job would entail. And I'm not going to say that much of it was incorrect. But I'm going to say that we

> have to keep before us, our children are able to perform. And that children will rise to the level of expectations. I was given the same negative press about the staff. The staff will rise to the level of expectations if they have the skills. And for the first quarter I was concentrating with my probation manager and my external partner on providing certain skills for our staff. We didn't do any evaluation, we did a lot of assisting, a lot of supporting to make sure that it was understood what was to be done, what we expected of them and what we expected for our children.

Having suspended evaluations, school leaders focused on building teacher capacity, using test results as opportunities to identify, mobilize, and develop the capacities thought necessary to improve student performance on the test. Therefore, accountability policies were perceived as an opportunity to build teachers' instructional capacity.

Probation status was also used to challenge and motivate teachers to change. The principal explained,

> When I looked at records, when I studied trends, when I looked at units and lesson plans, it . . . just felt that something different was needed and more desirable. This might not work either but we know that what we'd been, uh, in process for 4 years hadn't worked so, and that was my uh, logic to the staff. . . . And the staff has bought into all that has been presented. Sure there's some questions, there's some hesitancies but I don't think at this point they can afford not to because no one wants to be told if they had, so they're going to do it and if it doesn't they can say, "Well, we did it."

Dr. Sampson also reminded teachers repeatedly about her prerogative to replace teachers who were not pulling their weight, couching this prerogative as a "requirement" she would have to reluctantly enforce. In one faculty meeting, for example, she told teachers that she had been asked by the probation manager (and the central office) to begin the process of identifying people who might potentially be replaced at Waxton. She explained to the teachers that, although she "could no longer ignore" these requests by the probation manager, she would drag the process along to give each one of them "a chance" to improve.

Conclusion

Getting the school off probation in the 1st year of her tenure, the principal and her administrative colleagues managed to legitimate their approach with district administration and staff. Viewing accountability policy as a mechanism to motivate teachers and an opportunity to build instructional capacity among them, school leaders were able to press for change at Waxton. The challenge as the school moves forward will be to ensure that complacency does not set in with the removal of the external accountability pressures.

Moreover, although capacity building was a focus during the 1st year, there was a definite change in the principal's tone during the final meeting of the school year. At that meeting, the principal suggested that she would adopt a much tougher stance for the following fall. As the principal pushes for more internal accountability, and shifts her focus from support for teacher growth and development to expecting improved outcomes, it will be interesting to observe how teachers respond.

DISCUSSION AND CONCLUSION

Adopting a cognitive perspective on implementation, we used three cases to investigate the implementation of district accountability policies at the school level, paying particular attention to the manner in which school leaders enacted their environments and, in the process, mediated district accountability policies. More specifically, our account illuminates how school leaders, because of their position as middlemen between teachers and district offices, play the role of sense-maker with respect to accountability policies. However, rather than a straightforward response to high stakes or attractive rewards, school leaders' sense-making is influenced by the multiple overlapping contexts in which their work is nested. Therefore, schools' responses to district policies must be understood as a function not only of leaders' identities but also the multiple contexts in which their sense-making is situated. Answering to or enacting accountability policy meant something different, depending on the school. Viewing accountability policies from this perspective underscores how the implementation of district accountability policy has to be understood in terms of a two-way interaction in which accountability policy shapes and is shaped by the implementing agent and agency.

Making Sense of Accountability Policy: Similarities and Differences

There were similarities and differences among schools in the ways in which school leaders constructed district accountability policies. Beginning with the similarities, at least three patterns are evident. First, district student assessments and assessment data, especially the ITBS, figured prominently in all three schools. At Waxton, which faced the threat of school closure unless scores were raised, we expected to find and did find heavy emphasis on raising test scores. However, even in the two schools that had nothing to fear from district probation and reconstitution policies, tests and test data figured prominently for school leaders. Even at Dodge, where some teachers were not convinced that district accountability policies were salient considering their students' relatively successful performance on the ITBS, accountability policies figure prominently for school leaders. Leaders at the three schools paid careful attention to test score reports, describing probation and test-taking processes with great seriousness. However, how and why district accountability policies came to assume such importance for these schools was mediated in important ways by the wider professional and community contexts in which school leaders were situated.

A second and related pattern concerns the manner in which school leaders used test score data to gauge performance and think about their efforts to improve instruction. In all three schools we find evidence that the school administration translates test score shifts in terms of what teachers were doing in their classrooms, that is, the content they were covering and/or the ways in which they taught that content. District accountability policies and test score data in particular were being used by these school leaders to connect their leadership activities with instruction. This is important because most previous scholarship suggests that school management and teaching are "loosely coupled" (Meyer & Rowan, 1977; Weick, 1979). Our account suggests that in the implementation of district accountability policies, district leaders are attempting to connect their leadership work with teaching in tangible ways. Contrary to the institutional perspective, which suggests that policies are designed to keep constituents content and preserve institutional legitimacy, we find that school leaders enacted accountability policies in ways that sought to transform the core technology of the organization rather than simply buffer that technology from external scrutiny. Although accountability policies may serve to preserve the legitimacy of the institution and its leaders, these policies are also being understood and used as ways of leveraging change in instruction.

A third similarity was the manner in which district accountability policies focused school leaders' thinking and work around classroom instruction on literacy and, to a lesser extent, mathematics. In all three schools, we found that the talk and efforts of school administrations focused chiefly on literacy. The schools' emphasis on literacy and mathematics is not surprising, given that the school system had made improvements in these particular subject areas the primary measure of school progress, suggesting that district accountability policies and practices can have direct effects on schools' instructional priorities. However, in support of our larger argument, we also found evidence of how schools' sense-making about district priorities complicated their responses. Beneath more general subject-neutral statements such as "reaching benchmark levels" and "getting off probation," leaders framed and treated the test score challenge as a subject-specific issue. For example, the Waxton principal determined that literacy improvements were more pressing than mathematics—although the school scores were slipping in each and their removal from probation depended on improvements in both.

Although there were similarities across schools, implementing or enacting district accountability policies meant different things in these three schools. Our account, however, also illuminates distinct differences among schools. There were even substantial differences between the two schools that served similar student populations and had roughly similar student test score performance. At Waxton, as one might expect, district sanctions were front and center with the administration, and getting the school off probation through focused interventions that targeted the teaching of particular skills at particular grade levels was the central intervention. The approach might be best described as, in the words of Waxton's leaders, "assessment driven." Targeting grades that took the ITBS, the administration, together with its probation manager, developed a program of professional development for teachers that focused on the mathematics and reading skills covered by the test and test-taking skills. District accountability policies were a recipe for intervention at Waxton, used by the administration to define the content of capacity building efforts at the school as well as a stick that the administration held over teachers.

At Dodge, the school administration, especially Mrs. C., constructed the district's accountability policies as something of a blueprint for action, a means of improving the school's performance on standardized tests. However, the nature of the problem for Mrs. C and her colleagues was not so much about teaching particular skills that were on the test as it had to do with a combination of getting teachers to spend more time on task—teaching—and improving their instructional strategies. Still, cognizant of teachers' resistance, Dodge's administration was content to treat district accountability policies as optional and did not use the policies as a stick to leverage change with staff.

At Baxter, school leaders constructed and represented district accountability policies differently. Cognizant of teachers' critical stance toward the district's testing policy, the school administration (which had a strong commitment to site-based management and school-level autonomy) understood district accountability policies as a source of data that could be used by staff to analyze and document their progress and develop school improvement strategies. Using deliberate and careful longitudinal analysis of test data, Baxter's administration worked with teachers to identify problem areas in the curriculum and develop elaborate strategies to ameliorate these instructional problems. At Baxter, unlike Waxton, district accountability policies and test data in particular were not blueprints for instruction or school improvement but the starting point for deliberations among staff as to the nature of their problems and the development of workable solutions.

Schools varied then in the ways in which school leaders noticed and interpreted district accountability policies including test score data and district sanctions and rewards. These differences illuminate how the school-level mediation of district accountability policies transformed that policy in different ways, contributing to differences in the ways in which these mechanisms were represented to teachers.

Although accountability policies may define concrete outcomes with tangible rewards and sanctions, these accountability policies as enacted by school-level actors take on different forms and functions. Hence, accountability policy levers worked on and in schools in rather different ways, suggesting that the causal theories at work are much more complex than the rational choice and institutional models would lead us to believe.

Sense-Making as a Situated Activity

To understand these differences among schools in their enactment of district accountability policies, it is necessary to explore district leaders as sense-makers and especially how their sense-making is situated in multiple and overlapping contexts. Although district leaders in all three schools were situated in the same school district, there were important differences in who these leaders were and their situations.

To begin with, leaders at these three schools had rather different beliefs, histories, and agendas. Although the identity of school leaders rarely figures in discussion of accountability policy effects, we found these differences to be influential in leaders' policy responses. For example, the principal at Baxter was deeply committed to site-based management and involving teachers in making decisions about improving the school. These beliefs played a substantial role in the ways in which he constructed and implemented district accountability policies. Both the principal at Waxton and Dodge were new, working to establish their legitimacy with the district and peers as well as with their staff.

Situation played an important role in school leaders' sense-making. To begin with, the three schools were at rather different stages in the reform or change process. For schools like Waxton, district accountability policies were chiefly if not exclusively about getting off probation. At Baxter, a school that has enjoyed many successes during the past 5 years, district accountability policies became an occasion for developing curricula and instructional strategies that draw on the pooled expertise of Baxter teachers.

But the situation is more complex than where these schools were in the change process. School leaders were nested in multiple overlapping and interacting contexts that influenced and were influenced by their sense-making. One important context was students. Baxter and Dodge provide an excellent example of how the demographic characteristics of students were an important context in school leaders' and teachers' efforts to make sense of district accountability policies. Whereas Baxter and Dodge are located in the same neighborhood, serve roughly the same students, and have each experienced large increases in the percentage of low-income students in the past 10 years, these schools construct and respond to their students in very different ways. Baxter's leaders and staff constructed changing demographics as a challenge to be proactively addressed, whereas at Dodge these same changes provided a plausible explanation for declining test scores and more restrained change efforts. This example illuminates how social actors and schools as organizations infuse their environments with meaning, determining the implications of ostensibly "objective" environmental features.

Another important context in school leaders' efforts to make sense of district policy was teachers. At Baxter, teachers had a history of working together with leaders to make decisions with respect to school improvement and their teaching practice. The situation was very different at the other two schools. At Dodge, teachers worked on their own and antagonism rather than collaboration characterized their relationship with school leaders, especially the principal. The situation was similar although changing at Waxton. As the cases illuminated, these circumstances were especially influential when it came to district leaders' enactment of district accountability policy. For example, at Baxter, leaders used strong professional community and participatory structures to encourage teachers to deconstruct aggregated test score data. At Dodge, where such community was lacking, the principal introduced test score data to generate discussion, but the result was that it became an opportunity for faculty to resist school leaders and undermine district policy directions.

Principals' participation in other kinds of formal and informal professional networks also influenced their responses. In two schools, principals reported that their attention to test scores was partially motivated by their concerns for credibility in the eyes of neighboring schools and fellow principals. At the third school, Waxton, external partners (in this instance a neighboring school) influenced how the schools responded to district accountability mechanisms.

Middle-Level Managers as Intermediaries

One critical dimension in understanding how school leaders, especially the school principal, constructed accountability policies concerns their position as intermediaries between district office and teachers, intermediaries who depend on subordinates and superordinates for their success as school leaders. As was evident in all three cases, gaining and maintaining legitimacy with teachers as well as with peers and district office was an important part of the sense-making context for school leaders. Although their formal positions as leaders gave them some authority to define and interpret district accountability mechanisms for their staff, they had to gain the cooperation of their staff as well. This was important for all three principals, but it was especially important for the two newer principals who sought to establish their legitimacy with peers and their staffs.

Accountability mechanisms accentuated these tensions, pulling school leaders in two directions. District accountability measures, focusing on aggregations rather than individuals, pressed school leaders to think about their student body in the aggregate whereas teachers who thought more in terms of particular students pressed for more of a focus on the individual student. Institutional theories assume that faced with competing norms and pressures, principals do things that look like they are making improvements but do not cut deep below the surface. We found, however, that leaders' efforts to balance these tensions can have very concrete implications for the curricular priorities the school sets and teachers' classroom activities.

Managing in the middle presents rather particular dilemmas, especially when it comes to policies like accountability that threaten the autonomy that teachers cherish. Still, districts charge school leaders with enacting these unpopular policies with their staff while at the same time school leaders are dependent on teachers to get test scores up. Furthermore, principals are also caught trying to meet the demands of a policy system that operates on political time and that wants quick results, and classroom teachers, aware of the fads that plague the education policy arena, who want to introduce more incremental changes. Negotiating these two worlds, as school leaders must, is not easy.

Of course, managing in the middle in an era of accountability can also have advantages. Skillful school leaders can use accountability policies to augment their authority with respect to instruction. For example, the principal at Waxton used the threat of probation and firing teachers openly with her staff. Hence, school leaders can interpret district accountability policies in ways that support their own reform agendas and use them to augment their influence over staff. Because the stakes are high, they can use district accountability measures to add considerable clout to their own efforts to transform practice in particular ways.

NOTES

1. The following section draws heavily from Spillane and Reimer (in press).

2. While we will use the Consortium's data on "academic productivity," a weakness with this measure is that the Iowa Test of Basic Skills (ITBS) is inadequate to assess students' mastery of the more challenging reading and mathematics content. Furthermore, all of these measures are proxies for a schools' engagement in instructional improvement and improvement should not be attributed to school leadership.

3. Our interview questions were designed to get at five core issues about the practice of leadership:

   a. Getting the leaders to identify the key goals or macro functions they work on (e.g., building a school vision, promoting teacher professional development, improving test scores, etc.).
   b. Getting them to describe what day-to-day tasks they perform to attain these goals, that is, the micro tasks (e.g., observing classrooms, forming breakfast clubs, facilitating grade-level meetings, etc.).
   c. Getting them to describe how they enact the micro tasks, that is, their practice as leaders.
   d. Whether and how macro goals/functions and micro tasks are coenacted, that is, the extent to which their functions are executed with the help from others in the school.
   e. What tools and material resources (including designed artifacts, memos, protocols, organizational structures) the interviewees identified as important in the execution of macro and micro tasks.

4. Observation protocols focused on:

   - The nature and substance of the task—what the leader(s) did and the goals of the activities, including the subject matter focus of the activity, if any.
   - How the task was enacted, including the artifacts/materials used and how they were used to enable practice.
   - The timing and location of the task—the physical setting and context of the enactment, and the time of the year, week, or day on which the task was enacted.
   - The patterns of involvement, including what the leaders/facilitators did during the enactment, whether leadership was shared, and the role of participants.

REFERENCES

Blumer, H. (1969). *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.

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

Clotfelter, C., & Ladd, H. (1996). Recognizing and rewarding success in public schools. In H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 23-64). Washington, DC: The Brookings Institution.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality on organizational fields. *American Sociological Review*, *48*, 47-60.

Elmore, R., Abelmann, C., & Fuhrman, S. (1996). The new accountability in state education reform: From process to performance. In H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65-98). Washington, DC: The Brookings Institution.

Firestone, W. A. (1989). Using reform: Conceptualizing district initiative. *Educational Evaluation and Policy Analysis*, *11*(2), 151-165.

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. Berkeley: University of California Press.

Glaser, B. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory.* Mill Valley, CA: The Sociology Press.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Guthrie, J.W. (Ed.) (1990). *Educational Evaluation and Policy Analysis, 12*(3).

Hess, G. A. (2000). *Changes in students achievement in Illinois and Chicago, 1990-2000*. Washington, DC: The Brookings Institution.

Hill, H. C. (2001). Policy is not enough: Language and the interpretation of state standards. *American Educational Research Journal, 38*(2), 289-318.

Lin, A. C. (1998). *Doing my time: Policy implementation in prison*. Ann Arbor, MI, manuscript submitted for publication.

Lin, A. C. (2000). *Reform in the making: The implementation of social policy in prison*. Princeton, NJ: Princeton University Press.

Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public services*. New York: Russell Sage Foundation.

Lortie, D. C. (1975). *Schoolteacher.* Chicago: University of Chicago Press.

Meyer, J. W., & Rowan, B. (1977). The structure of educational organizations. In M. Meyer (Ed.), *Environments and organizations* (pp. 78-109). San Francisco: Jossey-Bass.

Miles, M., & Huberman, M. (1984). *Qualitative data analysis: A source book of new methods*. Beverly Hills, CA: Sage.

Resnick, L. (1991). Shared cognition: Thinking as social practice. In L. Resnick, J. Levine, & S. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 1-22). Washington, DC: American Psychological Association.

Scott, W. R. (1992). *Organizations: Rational, Natural, and Open Systems*. Englewood Cliffs, NJ: Prentice Hall.

Scott, W. R., & Meyer, J. W. (1991). The organization of societal sectors: Propositions and early evidence. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 108-142). Chicago: University of Chicago Press.

Spillane, J. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, *31*(2), 143-175.

Spillane, J. (2000). Cognition and policy implementation: District policy-makers and the reform of mathematics education. *Cognition and Instruction*, *18*(2), 141-179.

Spillane, J., Halverson, R., & Diamond, J. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher, 30*(3), 23-28.

Spillane, J., & Reimer, T. (forthcoming). *Implementation: A Cognitive Model*. Unpublished manuscript, Northwestern University, Evanston, IL.

Spillane, J., Reiser, B., & Reimer, T. (in press). Policy implementation: A cognitive model. *Review of Educational Research*.

Suchman, L. (1988). *Plans and situated actions: The problem of human/machine communication*. Cambridge, UK: Cambridge University press.

Weick, K. E. (1979). *The social psychology of organizing*. New York: McGraw-Hill.

Weick, K. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Wong, K., & Anagnostopoulos, D. (1998). Can integrated governance reconstruct teaching? Lessons learned from two low-performing Chicago high schools. *Educational Policy*, *12*(1-2), 31-47.

Yanow, D. (1996). *How does a policy mean?* Washington, DC: Georgetown University Press.