Where the Rubber Meets the Road: Rethinking the Connection between High-Stakes Testing Policy and Classroom Instruction

Author(s): John B. Diamond

Reviewed work(s):


Published by: American Sociological Association

Stable URL: http://www.jstor.org/stable/20452714

Accessed: 09/01/2013 15:39

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.
Where the Rubber Meets the Road: Rethinking the Connection Between High-Stakes Testing Policy and Classroom Instruction

John B. Diamond
Harvard University

In this article, the author examines the link between high-stakes testing policies and classroom instruction. Using data from classroom observations and interviews with teachers, he argues that these policies influence instruction but are mediated by teachers and filtered through their collegial interactions. He shows that teachers link the influence of high-stakes testing policies to instructional content (the knowledge and skills that they emphasize) more often than pedagogy (how they engage students around instructional content). As a result, didactic instruction dominates, especially in predominantly low-income and African American schools, in a policy environment that encourages addressing racial and class achievement gaps by increasing the use of interactive forms of instruction. The author concludes that researchers should be cautious not to overstate the impact of these policies on pedagogy and educational equity.

Understanding racial disparities in educational outcomes is a major concern in sociological research. Standards-based accountability policies are designed to challenge these inequalities by motivating teachers to expose all students to high-quality instruction along two critical dimensions: content and pedagogy (Lee and Wong 2004; Lipman 2004; Muller and Schiller 2000; Sandholtz, Ogawa, and Scribner 2004; Shulman 1986; Spillane et al. 2002). While these policies vary across state and local contexts (Rhoten et al. 2003), they generally emphasize transforming instructional content (the knowledge and skills that teachers teach) and pedagogy (how teachers engage students around instructional content).

Most prior work has suggested that high-stakes policies exert a major influence on instruction for better or for worse. Some research has found that these policies improve students’ outcomes by motivating educators to emphasize more rigorous content and by leading teachers to use pedagogical approaches that enhance students’ learning outcomes (Benveniste 1985; Bishop and Mane 1999; Bishop et al. 2001; Borko, Elliot, and Uchiyama 1999; Coleman et al. 1997; Kelly, Heneman, and Milanowski 2000; Shouse 1997; Wise 1979). Other work has indicated that accountability policies exacerbate inequalities by leading teachers to narrow the content they teach; marginalize low-performing students; or emphasize didactic pedagogy, characterized by lecture, seat work, memorization, and recitation—particularly in the lowest-performing schools (Anag-nostopoulos 2006; Booher-Jennings 2005; Clotfelter and Ladd 1996; Diamond and Spillane 2004; McDill, Natriello, and Pallas 1986; McNeil 2001; Spillane 2001).
However, given that educational reforms rarely lead to major transformations in instruction (Fullan 1991, 1993; Tyack and Cuban 1995) and that some have argued that the impact of accountability policy on instruction and equity may be limited (Darling-Hammond 1994, 2004; Firestone and Mayrowetz 2000; Firestone, Mayrowetz, and Fairman 1998; Lee and Wong 2004), it seems prudent to examine carefully how and to what extent accountability policies influence instruction.

In this article, I use data from classroom observations and interviews with elementary school teachers to examine how high-stakes testing policies influence instructional content and pedagogy in Chicago elementary schools. Building on theories of school organization and the faculty workplace, I first outline the multiple influences on teachers’ instructional decisions and place high-stakes testing policies in the context of this broader array of influences. I argue that policy messages are filtered through teachers’ interactions with their teaching colleagues and school administrators and are mediated by teachers themselves in ways that shape how these policies influence instruction. Second, I show that as a result of this process, teachers link the impact of accountability policy to the content they cover more often than to the pedagogical strategies they use. The fact that accountability policy influences teachers’ instructional content more than their pedagogy suggests that the strength of the link between the policy environment and the classroom varies across different instructional dimensions. Third, I explore the implications of the second finding for issues of equity. While Chicago’s accountability policy sought to transform both content and pedagogy in these schools, it seems to have had little impact on pedagogy, which many analysts believe is essential to reducing educational inequality.

BACKGROUND

Linking the Policy Environment and Instruction

The link between the policy environment and classroom instruction is a central issue in the sociology of education. Policy makers typically draw on a relatively straightforward model of organizational change—the bureaucratic/rational choice model. In this model, governmental agencies create policies that affect instruction by mobilizing rewards and sanctions that matter to school personnel (Spillane et al. 2002). This model suggests a strong link among the policy environment, the authority of school and district administrators, and classroom instruction. Although most scholarly work has acknowledged that mobilizing rewards and sanctions gets educators’ attention, the link between the policy environment and changes in instruction is complex, and studies in the neo-institutional and faculty workplace traditions have conceptualized this process differently from those that have used the bureaucratic/rational choice perspective.

Neo-institutional theory suggests that the policy environment and classroom practice are decoupled (or loosely coupled) (Bidwell 2001; Coburn 2004; Rowan 1990; Weick 1976). It indicates that because the core technology of schooling is uncertain and/or dynamic (Rowan 1990), the administrative apparatus of schools partially buffers classrooms from external inspection, helping to explain the failure of many reforms to penetrate classrooms fully.

With recent shifts in the policy environment—from a focus on compliance to an emphasis on accountability and students’ performance on standardized tests—some have argued that the connection between the policy environment and instructional practice has become tighter (Coburn 2004; Elmore, Abelman, and Furman 1996; Spillane and Burch 2006). However, while policy messages may penetrate classrooms, they are mediated by teachers who filter them through their prior practices and beliefs about teaching and learning (Coburn 2004:234; Spillane and Callahan 2000). For instance, Coburn (2004) systematically studied three teachers’ responses to reading policy in California between 1983 and 1999 and found that the teachers’ most common response to policy messages was “assimilation,” a process through which they interpreted policy messages on the basis of their prior practices and understandings about reading instruction.
Thus, although they changed “instructional routines, materials, or classroom organization” (Coburn 2004:224–25), they often altered “the underlying pedagogical or epistemological assumptions of the approach” they sought to implement (Coburn 2004:225). Hence, teachers’ interpretations of policy messages shape how they implement reforms in their classrooms and may influence which dimensions of instruction are most directly and meaningfully affected (Firestone and Mayrowetz 2000; Weiss et al. 2003).

While individual teachers mediate policy messages, broader organizational dynamics influence their beliefs and practices as well. Studies in the faculty workplace tradition have demonstrated that teachers’ instructional decisions are shaped by teaching colleagues through their interactions in informal communities of practice (Bidwell 2001; Booher-Jennings 2005; Lave 1988; Lave and Wenger 1991). Bidwell (2001:105) argued that teachers “turn to their colleagues for guidance and support, so that faculties become small, informal problem-solving social systems. Networks of colleague-to-colleague consultation and advice are the prime structural elements of these systems.” These interactions tend to influence teachers’ beliefs about instruction and how they teach (Bidwell and Yasumoto 1999). Likewise, other work has shown how school leaders’ interpretations of policy messages can influence how such messages are received and implemented in schools (Diamond and Spillane 2004; Spillane et al. 2002).

Therefore, rather than an unproblematic link between the policy environment and the classroom, some research has posited organizational and individual mediation of policy messages as they are introduced to schools and classrooms. Teachers draw on their prior experiences and beliefs, as well as their interactions with their teaching colleagues and administrators, as they interpret and implement policies. My data show that this process led teachers in the schools I studied to respond more fully to demands about changes in content as compared with pedagogy. I discuss the theoretical and practical implications of these findings in the conclusion.

**Standards-Based Reforms and Educational Inequality**

The recent accountability movement can be traced back to the 1983 report, *A Nation at Risk* (National Commission on Excellence in Education 1983), which argued that if the United States was to compete effectively in the emerging global economic order, it needed to train its young people for greater productivity in the workplace by raising academic standards. Thus, schools that were designed for industrial capitalism would have to be transformed to schools that prepared students for new forms of economic organization. Instruction needed to be transformed from a “factory model,” which emphasized obeying rules, memorization, and learning decontextualized knowledge, to instruction that focused on conceptual understanding, active engagement by students, and problem solving. Therefore, many contemporary accountability reforms emphasize exposing students to rigorous content but, perhaps more important, teaching that content in new ways that promote critical thinking, active learning, and problem solving.

Unfortunately, students of different races and social-class backgrounds have traditionally had different levels of access to high-quality instruction (Barr and Drebben 1983; Darling-Hammond 2005) and pedagogy emphasizing valued forms of knowledge (Anyon 1980; Apple 1979; Ayalon 1994; Bowles and Gintis 1976; Drebben and Gamoran 1986; Du Bois 1973). Students from socially advantaged groups (i.e., middle- and upper-class white students) often receive more demanding instruction that emphasizes critical thinking, problem solving, and active participation in learning (Barr and Dreben 1983; Clark and Peterson 1986; Gamoran 1986; Smith, Lee, and Newman 2001) and provides access to “social power and reward” (Anyon 1980:67). Other students (i.e., working-class and African American students) often receive instruction that is more practically oriented; involves more memorization and recitation; and prepares them for manual, clerical, or low-wage service-sector work.

Standards-based reforms are designed to
address educational inequality by exposing all students to high-quality instruction and eliminating the differentiated pedagogy that many have argued undermines the achievement of low-income students and students of color (Sandholtz et al. 2004; Spillane 2001). While differences in instruction are not the only issue that contributes to educational inequality, the current emphasis on more rigorous instructional content and pedagogy emphasizes exposing all children to the kind of instruction that was once reserved for children of the social and economic elite.

**Interactive versus Didactic Pedagogy**

The distinction between interactive versus didactic pedagogy has been central to the discussion of race, class, and educational inequality. Middle-class white students are more likely to participate in interactive pedagogy, while low-income students and students of color are more likely to be exposed to didactic pedagogy. Smith et al. (2001) wrote that the core features of didactic instruction include teacher-led instruction, which emphasizes lectures; questions leading to “single, short answers”; and teachers’ assessments of the correctness or incorrectness of students’ responses. Students “listen to teachers and recite answers” and “try to repeat the knowledge they have been taught as it was transmitted” (Smith et al. 2001:11). Didactic instruction is similar to conventional teaching, which is “organized through a set pattern of lecture, recitation, and seat work (Gamoran, Secada, and Marrett 2000:57; Goodlad 1984).

In contrast, interactive instruction involves teachers coaching and guiding students through their learning, seeking explanations from students for their answers, and assessing the thinking that led to the students’ answers. Students discuss and ask questions about classroom work with the teacher, as well as with other students, and attempt to connect knowledge to prior understanding. Interactive (as opposed to didactic) instruction enhances gains in learning for students from all backgrounds as measured by standardized test scores. Smith et al. (2001:10) argued that “prior research has documented substantial achievement benefits, and no consistent disadvantages, for students exposed to . . . interactive instruction” (see also Knapp, Shields, and Trumbull 1992; Tharp 1982). In addition, the less standardized structure of interactive classrooms may lead to less stratified and stagnant conceptions of academic ability among students (Rosenholtz and Wilson 1980; Simpson and Rosenholtz 1986).

Unfortunately, data from a survey of teachers in Chicago schools during the 1996–97 academic year showed that low-income and African American students were the least likely to be exposed to interactive instruction (Smith et al. 2001). While state standards vary, interactive instruction is the form of teaching and learning that is most often associated with the goals of standards-based reforms (and seemingly supported by the Illinois standards). The interactive model not only enhances achievement, but provides more access to privilege in adulthood by giving students greater exposure to autonomy, self-direction, higher-order thinking skills, problem-solving ability, and complex communication skills (Anyon 1981). Such skills are associated with higher-status (and higher-paying) jobs (Levy and Murnane 2004).

In analyzing my classroom observation data, I characterized pedagogy in terms of interactive and didactic instruction and examined how the racial composition of schools was associated with these patterns within Chicago’s standards-based environment. My findings show that teachers more frequently respond to demands for changes in content than to changes in pedagogy. The stagnancy of pedagogy in these schools meant that classrooms in predominantly African American and low-income schools were dominated by didactic instruction.

**STATE AND LOCAL POLICY CONTEXTS**

**The Illinois Policy Context**

The Illinois Learning Standards (Illinois State Board of Education 1997) were established in 1997. The philosophy behind these stan-
standards maps onto much of the broader standards movement, as evidenced by the follow-
ing passages from the Illinois State Learning Goals. A portion of the introductory state-
ment reads:

Technological breakthroughs, an explosion of information and global economies are just a
few of the conditions that have changed dramatically in the past decade. To be successful
in a world characterized by change, students need to learn the basics, but the basics of the
1990s and the new century to come go far beyond the basics of the 1960s, 1970s or
1980s. In addition to basic knowledge, stu-
dents need to acquire new ways to learn that
will serve them throughout their lives.

Specific examples of these “new ways to
learn” are discussed in each subject matter
area. For example, in mathematics, students
are expected to be able to solve problems,
communicate about mathematics, work on
teams, and make connections across learning areas. In solving problems, students should
be able to “recognize and investigate prob-
lems; formulate and propose solutions sup-
ported by reason and evidence.” In commu-
nicating in mathematics,

students must have opportunities in mathe-
matics classes to confront problems requiring
them to translate between representations,
both within mathematics and between math-
ematics and other areas; to communicate
findings both orally and in writing; and to
develop displays illustrating the relationships
they have observed or constructed.

Finally, students should be given the chance
to work on teams and, in team problem solv-
ing, “to be prepared to function as members
of society and productive participants in the
workforce.”

These goals focus on what students should
know and be able to do, emphasizing that
“the same high standards should apply to all
students.” While the origins, politics, and
strategies of accountability policies vary
across states (Conroy, Elmore, and Siskin
2003; McNeil 2001; Rhoten et al. 2003), the
Illinois standards suggest an implicit push for
interactive instruction. Solving problems,
communicating about mathematics, working
in teams, and making connections across
content areas suggest instruction that pushes
students to think deeply about their work,
engage in conversations with their classmates
about what they are learning, and communi-
cate the reasoning that leads to their inter-
pretations of classroom materials.

The Chicago Policy Context

Chicago provides a fruitful context for study-
ing standards-based accountability policy. Its
accountability policy occurred in two differ-
ent ways during two historical periods. The
first phase, 1989–95, began when the
Chicago School Reform Act (P.A. 85-1418)
decentralized decision making and gave sub-
stantial power over budgetary decisions and
the hiring and retention of principals to par-
ents and community members. The 1995
Chicago School Reform Amendatory Act
transferred significant power to an appointee
of the mayor, the CEO, and empowered that
person to place failing schools on academic
probation and in remediation using results
from the Iowa Test of Basic Skills (ITBS)2 as the
indicator of a school’s success or failure. In
1996, 25 percent of the district’s elementary
schools (109 schools) were placed on proba-
tion (Hess 2000; Wong and Anagnostopoulos
1998). The vast majority of these schools
were low income and predominantly African
American. In fact, “of the 147 elementary
schools placed on probation from 1996-
2001, 75% were predominantly African
American schools” (Bryk 2003:255). Given
these data, some have argued that, in its ex-
cution, the accountability policy after 1995
emphasized using sanctions to improve the
performance of low-income African American
schools (Bryk 2003). The Chicago reform
placed most of its emphasis on sanctions for
schools and students. More than 50,000 stu-
dents who had not met grade-level perfor-
manve requirements were forced to partici-
pate in summer school during the first two
years of the program (Bryk 2003).

An important part of this accountability
process was the alignment of the State Goals
for Learning, the Chicago Academic
Frameworks, and the Chicago Curricular
Framework Statements. The Chicago
Academic Frameworks, as required by state
law, were mandated to "meet or exceed goals established by the state" (Chicago Public Schools 2000b). However, the Chicago policy has been criticized for its scripted curriculum and emphasis on the ITBS. In some sense, both this curriculum and the ITBS (with its emphasis on basic skills) may have sent mixed messages to Chicago teachers because while the Illinois standards emphasized critical thinking and problem solving, the scripted curriculum and ITBS sent signals that leaned toward instruction in basic skills. Nonetheless, the espoused intention of the Chicago policy was to provide "a strong focus on cultural diversity, critical thinking, problem solving, and decision-making skills" (Chicago Public Schools 2000b).

Since the inception of the policy, students' performance on the ITBS has improved substantially. The proportion of students who performed at or above the national norms in mathematics rose from 37 percent in 1997 to 46.6 percent in 2005, and the proportion of students who performed at or above the national norms in reading increased from 35.4 percent in 1997 to 43.7 percent in 2005. However, it is unclear if these changes represented real learning gains or inflated test scores (Bryk 2003; Jacob 2003; Koretz 2005). Moreover, with a few notable exceptions (Lipman 2004), not much is known from prior research about how, and to what extent, these policies shaped instruction in Chicago classrooms. This article begins to fill this gap in knowledge.

**METHODS**

The study reported here drew on data from the Distributed Leadership Project, a multi-year study of leadership practices in Chicago elementary schools (Spillane 2006; Spillane and Diamond 2007). This article relies on data collected in eight case-study sites during the 1999–2000 school year, the first year of the research. Table 1 shows the schools' demographic characteristics.

<table>
<thead>
<tr>
<th>Schoola</th>
<th>Student Enrollment</th>
<th>Low Income (%)</th>
<th>Black (%)</th>
<th>White (%)</th>
<th>Latino/Latina (%)</th>
<th>Asian (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>School 1</td>
<td>750–1,000</td>
<td>90–100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>School 2</td>
<td>1,000–1,500</td>
<td>60–70</td>
<td>&lt; 10</td>
<td>40–50</td>
<td>20–30</td>
<td>20–30</td>
</tr>
<tr>
<td>School 3</td>
<td>1,000–1,500</td>
<td>70–80</td>
<td>&lt; 10</td>
<td>40–50</td>
<td>10–20</td>
<td>30–40</td>
</tr>
<tr>
<td>School 4</td>
<td>250–500</td>
<td>90–100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>School 5</td>
<td>750–1,000</td>
<td>90–100</td>
<td>&lt; 10</td>
<td>0</td>
<td>90–100</td>
<td>0</td>
</tr>
<tr>
<td>School 6</td>
<td>250–500</td>
<td>90–100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>School 7</td>
<td>1,000–1,500</td>
<td>90–100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>School 8</td>
<td>1,000–1,500</td>
<td>90–100</td>
<td>&lt; 10</td>
<td>&lt; 10</td>
<td>80–90</td>
<td>&lt; 10</td>
</tr>
<tr>
<td>School 9</td>
<td>500–750</td>
<td>60–70</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>School 10</td>
<td>500–750</td>
<td>80–90</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>School 11</td>
<td>1,000–1,500</td>
<td>80–90</td>
<td>&lt; 10</td>
<td>20–30</td>
<td>20–30</td>
<td>50–60</td>
</tr>
<tr>
<td>School 12</td>
<td>500–750</td>
<td>60–70</td>
<td>&lt; 10</td>
<td>50–60</td>
<td>40–50</td>
<td>&lt; 10</td>
</tr>
<tr>
<td>School 13</td>
<td>750–1,000</td>
<td>90–100</td>
<td>20–30</td>
<td>&lt; 10</td>
<td>70–80</td>
<td>0</td>
</tr>
</tbody>
</table>

a Schools 1–8 are the case study schools that are the focus of this report. The other 5 schools (9–13) were interview-only sites. Demographic ranges are provided to protect the confidentiality of the schools and participants.

b The percentage of students who were eligible to receive free or reduced-price lunches serves as a proxy for school social class composition.
**Data Collection**

During the first year of the study, my colleagues and I spent 50–70 days observing formal and informal settings in the eight schools, interviewing teachers and administrators, and shadowing school leaders during the school day. In addition, we observed 105 predominantly second- and fifth-grade classroom lessons (involving 47 different teachers) (see Table 2 for detailed information on the observations) using a standard classroom observation protocol (for excerpts of the protocols, see the Appendix). Each lesson in these schools was approximately 50 minutes long. Building on the work of Doyle (1983), we focused on the academic tasks that typically make up any lesson (e.g., a whole-group discussion analyzing a story) and paid attention to whole-group, small-group, and individual students’ tasks in our observations. We also documented the amount of time spent on each task and what happened during transitions between tasks.

We focused on the content (the knowledge, concepts, and skills being conveyed, such as subtracting fractions) and the pedagogy of the tasks (strategies and mechanisms through which content is conveyed, including the teacher’s lecture and class discussion). In capturing these teaching strategies, we focused on discourse patterns, grouping arrangements, and the materials used throughout the lesson. After creating a narrative summary of classroom observations, we completed a set of closed-ended items. These items emphasized the pedagogy related to the academic tasks (and the content to a lesser extent) and included sections on students’ grouping arrangements and classroom discourse (for relevant items from the protocol, see the Appendix). We characterized entire lessons by answering questions such as this: “In this classroom during reading lessons, most questions were asked (a) only by the teacher, (b) mostly by the teacher, or (c) by teachers and students equally.”

These classroom observations were preceded and followed by interviews with the teachers. The preobservation interviews asked the teachers about the content and purpose of the lesson to be observed and if any aspects of the lesson deserved special attention from the researchers. The postobservation interviews focused on ascertainment the teachers’ perceptions of instructional changes they had made and the influences on these changes. More specifically, we asked the teachers about the content they were covering during the lesson and the specific pedagogy that we had observed during the class (e.g., students engaging in group work or students explaining their mathematics calculations on the chalkboard). The teachers were then asked how they had decided on their instructional strategies (e.g., “Why do you do this?”), if these pedagogic practices represented changes from their past instructional approaches (“Have you always [done this] in your teaching?” “If not, how long have you done it this way?”), and if they were influenced by anyone or anything in making these changes (“Did anyone or anything contribute to this change?”). If the teachers identified influences, they were asked to go into more detail about the nature of the influ-

<table>
<thead>
<tr>
<th>School</th>
<th>Number of Observations</th>
<th>Number of Teachers</th>
<th>Number of Multiple Teachers</th>
<th>Number of Language Observations</th>
<th>Number of Math Observations</th>
<th>Number of Science Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>8</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>19</td>
<td>6</td>
<td>5</td>
<td>10</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>105</td>
<td>47</td>
<td>28</td>
<td>42</td>
<td>43</td>
<td>20</td>
</tr>
</tbody>
</table>
ences. In this way, we were able to capture teachers’ perceptions of the influences on their instructional practices.

In cases in which we had not observed the teachers, we asked similar questions about instructional changes. If the teachers had made changes, we followed a similar line of questions as the one just outlined. We asked nonobserved teachers who reported making no instructional changes to imagine what they would do if they were to make changes in their practices and then followed a similar line of questions as the one just outlined.

Data Analysis

Having observed the teachers and asked them about specific instructional decisions and influences on them, we gained substantial information about the constellation of influences on instruction as perceived by the teachers. We first coded all the interview data using the topic code (Richards 2005) “influences.” We then proceeded to code the interview data focusing on who (principals or other teachers) or what (testing or standards) influenced the teachers and the focus of these influences. The coding for “influences” was inclusive, meaning that teachers could report multiple influences for any given instructional decision. Emerging from the teachers’ reports, the code for “who” included other teachers, principals, assistant principals, curriculum coordinators, and students, while the code for “what” included testing, standards, textbooks, and the Internet (among a few others). We further coded the data around the “focus” of the influence. For instance, if the teachers reported that they were influenced by the principal, we also noted the focus of that influence. Was it focused on a certain subject area (mathematics, science, or language arts instruction)? Was it focused on a certain dimension of instruction (content or pedagogy)? Therefore, we captured the source and instructional target of each influence. We also analyzed the interactions across influences. We were particularly interested in understanding if other teachers and administrators shaped the teachers’ efforts to respond to accountability policy. Therefore, in addition to using an inclusive coding process, in which multiple influences on each instructional decision were coded, we analyzed each case in which the teachers reported being influenced by administrators and other teachers to determine if these influences were directly associated with accountability policy.

The classroom narrative summaries were reviewed in relation to the same issues as the closed-ended items, emphasizing classroom grouping arrangements, discourse patterns, and the nature of the academic tasks. The analysis of interviews and observations was assisted by the computer program Nu*DIST 4.0 (and later the upgraded N6 program). I used the text and node searching capabilities of the program to identify patterns of influence across the data.

For the analysis reported here, I focus on teachers’ reports of the influences on instruction that derive from testing and standards. I argue that testing and standards are both part of a high-stakes testing policy regime because the Chicago Public Schools (2000a) sought to align its local standards, curricular frameworks, and testing with the Illinois State Board of Education’s learning goals. I report data from 84 teachers who discussed specific instructional practices that were tied to multiple influences, including testing policies and standards.

As I analyzed the data, I focused on issues of content and pedagogy because these are two of the important targets of high-stakes testing policy. I examine teachers’ reports of how testing and standards influenced instruction and distinguished between the impacts of testing policy on content versus pedagogy. When I found that testing policies had limited impacts on pedagogy, according to the teachers, I examined how the teachers’ pedagogy varied on a continuum between didactic and interactive instruction across schools that served students from different racial and social-class groups.

FINDINGS

The Multiple Influences on Instruction in General

Most previous research has suggested that high-stakes testing policies exert a powerful influence on classroom instruction. While
these policies do influence teachers’ instructional decisions, my data show that they are one among many influences on teachers. When we asked the teachers about specific recent changes in their instructional practices, many influences emerged. For instance, 83 percent of the 84 teachers indicated that principals influenced their instruction, 28.6 percent mentioned assistant principals, and 79.8 percent identified other teachers. Likewise 46 percent of the teachers reported being influenced by standards, and 42 percent reported being influenced by testing (also reported in Spillane, Hallett, and Diamond 2003). However, other nonhuman influences, such as textbooks and the Internet, were also influential (although to a lesser extent). A first-grade teacher discussed the multiple influences on her efforts to change her instructional practices when she transitioned from second- to first-grade teaching:

Over the past two years, I’ve changed tremendously in the way that I teach reading. So I had to pull things out of everywhere. I went to every teacher that I could possibly think of in this school that could help me. I got information from the reading recovery teachers, from the reading specialists. I went to the library a lot. I did a lot of reading on professional books. I was on the Internet all the time. I joined a first-grade newsletter on the Internet where people share different ideas and how to teach different things, and I tried everything.

This teacher’s instructional decisions, like those of the vast majority of the participants in this study, were influenced by multiple factors, including other teachers, school leaders, and nonhuman influences like resource guides and the Internet. Another teacher discussed the influence of textbooks and district frameworks on her instructional decisions following an observed classroom lesson: “What we did today was out of a fourth-grade book. The [Chicago] public schools also provide . . . like a structured curriculum book. We might pull some stuff from that; we kind of use that as a guide.”

Therefore, when it came to these teachers’ general instructional decisions, testing and standards were influential, but the teachers perceived the direct influence to be less powerful than their interactions with their teaching colleagues and school administrators. However, some of the teachers’ interactions with their teaching colleagues and administrators were efforts to respond to accountability policies. When I analyzed occasions in which the teachers were influenced by other teachers and administrators, I found that these influences were associated with testing and standards in 22.6 percent of the cases. Therefore, in a significant minority of cases, the teachers’ interactions with colleagues were in response to accountability policy. This finding suggests that accountability policy exerts a direct effect on teachers’ practices and, at times, an indirect effect as it is filtered through the teachers’ interactions with their colleagues.

The Influence of Testing on Instructional Content

The Illinois standards seek to affect instructional content (the knowledge and skills that teachers teach) and pedagogy (how teachers teach that content). To understand the impact of testing on instruction, I analyzed teachers’ reports of influences on specific instructional decisions within both dimensions. In this section, I discuss teachers’ reports of the influences on instructional content.

Figure 1 depicts teachers’ reports of the human and material influences on the instructional content that the teachers covered. With regard to content coverage, the most frequent influences were other teachers, the teachers themselves (self), textbooks, and standards. Of the 84 teachers who were interviewed, 56 percent said that other teachers influenced the instructional content they covered, 55 percent were influenced by textbooks, 51 percent relied on their own experience, and 48 percent were influenced by standards. Principals, students, and testing also emerged as relatively important influences on the content that the teachers covered. However, the teachers noted that they did not shape instructional content as often. Principals influenced 26 percent of the teachers; students, 26 percent; and testing, 31 percent. Thus, testing and standards shape the instructional content that teachers cover.
Diamond (Booher-Jennings 2005; Coburn 2004), which suggests a relatively powerful link between the policy environment and this instructional dimension.

How does high-stakes testing policy influence the content that teachers cover? About half the teachers we observed and interviewed argued that testing and standards had a direct impact on the content they covered. As one teacher stated, “Everything that I . . . choose comes from their standards, comes from what the Board of Education mandates that I teach and what [the students] will be tested on in May on the Iowas.”

The teachers whose content coverage was influenced by testing and standards identified four primary influences. First, as a result of testing, they focused on mathematics and language arts instruction more than on science and social studies instruction. They reported teaching science and social studies intermittently or after the testing period when the material in mathematics and language arts had been covered. Second, testing and standards led the teachers to focus on specific issues within subjects. Third, the teachers covered material more quickly to complete it before students were tested. Fourth, many teachers reported spending time on test preparation through the use of practice tests and the simulation of the testing situation during the regular school day. I briefly discuss each of these four influences next.

**Emphasis on Mathematics and Language Arts**

Research has demonstrated that teachers tend to neglect certain subject areas at least in partial response to high-stakes testing policies. For instance, science and social studies instruction is sometimes given less attention than is mathematics and language arts instruction (Spillane et al. 2002; Wills 2006). The teachers in this study reported a similar pattern. First, they focused most of their attention on mathematics and language arts instruction (the subject areas that were the

![Figure 1. Teachers' Reports of the Influences on Their Coverage of Classroom Content (based on the reports of 84 teachers)](image-url)
core focus of the district’s accountability policy. A first-grade teacher emphasized this instructional focus in an interview. “Well, our principal says . . . if you have to skip everything else, that is fine as long as you get math and the reading done. Those [math and reading] are the two things that [the students] are tested on.” In this case, the focus on mathematics and language arts instruction is influenced by the school principals’ interpretation of the importance of these subjects in response to pressures for accountability. This finding demonstrates how the impact of the policy environment can flow to teachers indirectly through their interactions with school leaders.

Science was consistently given less attention according to the teachers. For example, one representative teacher indicated that although students receive instruction in science and social studies, this instruction is intermittent in the lower grades and increases as these subjects are more prominently featured on tests in the upper grades. The teacher argued that “the seventh graders focus on science and social studies because that’s where they test.” Therefore, students at certain grade levels are more likely to get science and social studies instruction, while students at other grade levels receive less of it. Other teachers said that they delayed science instruction until after testing was completed. A third-grade teacher said that following testing, “we’re able to breathe, and all the things that you wanna do with the children you’re able to. . . . I had a chance to do science. . . . I had to wait ‘til after testing. . . . In third grade, they don’t test for science or anything, only reading and basically math.”

Therefore, in her third-grade classroom, this teacher narrowed the content she taught to the subjects that could affect the testing outcomes of her students. The third grade is one of the benchmark grades in Chicago in which students who do not reach a certain threshold must attend summer school and, if their scores do not meet this threshold at the end of summer school, they must repeat the third grade. Therefore, there is added pressure at this grade level for teachers to prepare students for the test.

Instructional Focus in Mathematics and Language Arts

Within subject matter areas, the teachers used testing data and the material that they anticipated would be covered on the tests to identify the particular instructional content on which they needed to focus. For instance, several teachers discussed the need to teach vocabulary to students (typically entire classrooms and grade levels) who had not performed well in this area on past tests. A second-grade teacher said, “I inundate my students with vocabulary because their vocabulary is so poor . . . and in order for them to be successful on the Iowa Test, they must be familiar with these words.” In this case, the teachers’ response to testing policy was to give students specific skills with which to succeed in particular areas of the test. Another second-grade teacher explained, “I [was] encouraged to really beef up the vocabulary because . . . when we take the Iowa test, that’s usually the lowest score for our kids.” This teacher indicated that she was “encouraged” to emphasize vocabulary in her classrooms by others in the school, again suggesting that some of the influences that come from other school personnel are related to the accountability policy.

Here, the instructional content is heavily driven by the ITBS. While some would argue that this type of reallocation of instructional time is appropriate, focusing on specific tested material can lead to the inflation of test scores without enhancing students’ learning (Koretz 2005). For instance, students may learn a set of words on which they are likely to be tested without expanding their knowledge of the broader domain of vocabulary words (Koretz 2005).

As was the case with language arts instruction, the teachers also focused on subdimensions of instruction in mathematics on the basis of the testing data. A fifth-grade teacher said the following when asked about an observed lesson: “We were just going over . . . subtraction of fractions . . . because I know they have a lot of those on the test; . . . fractions are always kind of a problem with kids.”

Another second-grade teacher discussed her decision to focus on problem solving in
her second-grade classroom. “I noticed on the Iowa test some of them still had difficulty with problem solving—especially multiple-step problems. So that’s an area that I’m really trying to bone up on.” For these teachers, instruction was focused on the specific content to be covered on the test and the teachers’ experiences with students’ prior outcomes on these tests. These responses potentially narrowed the content being covered and led to a focus on the specific demands of the ITBS.

While content is influenced by high-stakes testing policies, some teachers discussed mediating the impact of standards and testing. The second-grade teacher just discussed talked about the influence of standards on the instructional content that she covered in mathematics, language arts, and science. Following a classroom observation, she explained the connection of the material she taught in science that day to standards. “The science was state goal 12, academic standard B, frameworks 4–7, and then I can go to the book and look that up.” While this seems to be a powerful external influence on her choice of content, she added:

Frameworks 4–7 describe relationships among various organisms in their environment, describe ways in which organisms cause changes in their environment, and describe characteristics of plants and animals that allow them to live in specific environments. It all falls into place. It may not say specifically the desert habitat [which was the focus of her lesson], but you have to hone in and find where it fits.

In this case, while the standards and frameworks guided the material the teacher covered, she chose the focus of her instruction. She could have just as easily chosen an ocean, forest, or mountain habitat as the location to explore these issues. Her response to accountability policy also suggests a reciprocal relationship in which the broad parameters of the content she covered were shaped by the policy environment, but she also mediated what that meant in practice by finding where the standards “fit” in her instructional plans and arguing that “it all falls into place.” This statement suggests a certain degree of autonomy and that the teacher was “assimilating” policy pressures into her preexisting practices.

A fifth-grade teacher at another school stated, “I definitely try to teach those goals and skills that are required, but after I have met those, or within those, I do what I feel the children should know.” These comments suggest that even when external influences are powerful (when teachers explicitly connect what they teach to standards and testing), some teachers mediate the linkage between the policies and what they actually teach. Some teachers respond to the general demands of the standards but select the content they see as appropriate for their specific lessons. In addition, the fifth-grade teacher just quoted supplemented teaching the “skills that are required” by teaching important information that she “feels the children should know.” This process is similar to the “bounded autonomy” discussed by Coburn (2004). Some of these teachers made instructional decisions within the parameters set by the accountability system; however, their decisions were not completely out of their hands.

**Sequencing and Pacing of Instruction**

Testing policy influenced when during the school year and at what pace the teachers taught material. In many cases, the teachers reported fitting material in prior to testing that would normally come after testing. As one second-grade teacher explained, “There [are] some math topics that I would normally let the year progress I would not have to cram in. . . . I have to do a quick minilesson about it.” Other teachers reported “cramming” material in prior to testing or pushing forward on course content when their students had not fully mastered it.

**Test Preparation**

A final point on content is that many teachers reported spending time on preparing the students for tests, including covering tested content and coaching students on test-taking skills. Thus, some content was missed to prepare students for the testing situation. The
teachers in every school reported engaging in test-preparation activities. Some schools engaged in substantial test preparation. The following comments are from teachers at two schools:

Every week we have, beginning the first week of school, an hour practice testing every Thursday morning. In addition to that, I believe that the teachers probably spend, you know, another hour during the week reviewing just the testing and then when the testing time comes, for about a month before the test every day. (Special education teacher)

Like, toward more like March [or] April, we do a lot of timed tests because [the tests are] timed. I have to give them practice with that . . . you know, where you have to go over the tests prep books. . . . It’s not fun, but . . . for their success, I have to do that. (Second-grade teacher)

Although test preparation was substantial, far less time was spent on it than on regular classroom instruction. Even in the school that spent the most time on test preparation, these activities were limited to one to two hours per week during the regular school year and increased to at most five hours per week in the month before testing. Thus, while the teachers were drilling certain content and using practice tests (provided by the test manufacturers and containing similar content) to prepare their students for testing, these activities were time limited and did not seem to affect the teachers’ pedagogical practices during typical class lessons.

With regard to the content that the teachers covered, these data show that the teachers reallocated (Koretz 2005) instructional time to certain subject areas (mathematics and language arts) and emphasized specific issues within subjects. Some teachers also reported that the sequencing and pacing of instruction were influenced by testing and that instructional time that could be spent on other topics was spent on test-preparation activities. About half the teachers said that testing and standards directly shaped the content they covered. Other teachers’ comments suggest that the influences of high-stakes testing may be more indirect, with teachers being “encouraged” or directed to cover certain content by school leaders who are driven to respond to accountability demands. This finding suggests that testing influences instructional content for a large percentage of teachers but does not completely dominate all teachers’ decision making. As some of the teachers quoted earlier indicated, even when they responded to the demands of accountability policy, they also had some discretion. For instance, while the policy may specify certain content to cover, the teachers reported extending this content to include other important material that they thought the students should know and fitting such content demands into existing priorities and practices.

Whether teachers are being focused on more important content by such policies (as supporters would suggest) or are facing a dangerous narrowing of the content they teach (as opponents would argue) neither process seems to dominate instruction completely in these classrooms. In fact, many teachers believe that testing and standards do not influence the content that they cover, and many others think that they have some discretion even within the constraints of the high-stakes system.

The Influence of Testing on Pedagogy

The teachers’ reports indicated that testing and standards were more powerful influences on content than on pedagogy. As Figure 2 shows, the teachers reported that they (63 percent) and their teaching colleagues (70 percent) were the most important influences on their pedagogy. For instance, a fifth-grade teacher discussed seeking guidance on language arts instruction from her colleague, Mrs. Diaz, who “is a strong language arts person. . . . Whenever . . . I . . . want to know about how to go about a strategy a particular way, I might ask her, “Well, how do you do this?” and “Does this work well with your students?” The teachers often sought guidance through such informal interactions. Students (44 percent) and textbooks (40 percent) were also important influences on teachers’ pedagogical decisions. In contrast to their relatively strong influence on content, however, the
teachers did not view testing (19 percent) and standards (18 percent) as major influences on their pedagogical decisions.

Thus, when deciding how to teach, the teachers reported being influenced most heavily by their own thinking, their teaching colleagues, their students, and textbooks and far less by standards and testing. One representative teacher described the pedagogical limitations of the standards. “It just tells you what you have to teach, like antonyms, synonyms, homonyms, prefixes, suffixes, compound words, contractions. It doesn’t tell you how to teach it.” She continued to describe her pedagogical influences: “Many of those strategies come from either my own experience from being a teacher, where my students are, different classes that I’ve taken, methods classes that teach you how to teach certain subjects, different places. Teacher’s editions, real-life experience.”

The difference in the influence of testing and standards across content and pedagogy was striking. While 48 percent of the teachers thought that standards influenced the content that they taught and 31 percent thought that testing did so, only 18 percent thought that standards influenced their pedagogy and 19 percent thought that testing did so. This finding suggests that the extent to which the policy environment penetrates classrooms depends on the instructional dimension being examined. The data suggest that standards and testing alone are not overwhelming influences on pedagogy in these classrooms. When teachers are making pedagogical decisions, they are most likely to go to other teachers for guidance, reflect on their own experience, look to textbooks for ideas, or reflect on their students’ learning styles and perceived capabilities. This finding is in line with research on the importance of collegial interaction in influencing teachers’ instruction (Bidwell 2001). Therefore, the direct impact of high-stakes testing policy on pedagogy in these schools seems limited.

Among teachers who reported that their pedagogy was influenced by testing and stan-

Figure 2. Teachers’ Reports of the Influences on Their Teaching Strategies
Where the Rubber Meets the Road

299
dards, responses varied. Some argued that they sought to push students toward higher-order thinking skills, to explain their answers in mathematics, or to work in cooperative teams. A comparable number of teachers focused more on basic skills, recitation of correct answers, and whole-class instructional methods. A few expressed tension regarding how to respond simultaneously to the demands of the testing system and of school leaders. During a postobservation interview, one teacher discussed the difficulty in deciding on instructional strategies in mathematics. Spearheaded by the principal, her school had adopted a progressive approach to mathematics instruction, which led the teachers to be frustrated because “it doesn’t give the kids enough practice or . . . enough background” to be successful on the ITBS.

This teacher’s comments demonstrate the ambiguity that is sometimes built into the policy environment. While the standards are in line with this mathematics curriculum, the ITBS is not (at least as far as this teacher was concerned). Moreover, the response to these external demands was mediated by the teacher’s beliefs about her students’ capabilities, the influence of the school’s principal on her instruction, and the curriculum that the school had chosen. The teacher’s comments (and the varied responses of other teachers) demonstrate that definitive claims about the impact of these policies on instruction may be misguided. While proponents and opponents of testing policies argue that they will result in clearly defined responses that are in line with the spirit of these policies or out of line with them, my data demonstrate that the direct influence of testing on pedagogy is limited to a small number of teachers and varies among them.

The Nature of Pedagogy and the Testing Environment

One implicit goal of the high-stakes reform was to transform pedagogy so that interactive instruction would be more common. Such instruction would include students solving problems, engaging in teamwork, and communicating the reasoning behind their mathematical calculations. This interactive pedagogy suggests instruction that encourages students to think deeply about their work, discuss and ask questions about classroom material, interact with their classmates about substantive issues, and communicate the reasoning that leads to their interpretations of classroom materials. Research conducted near the policy’s inception (1996) showed that while all students tended to perform better on the ITBS when they were exposed to interactive instruction, students in high-poverty and predominantly African American schools received more didactic instruction (Smith et al. 2001).

To examine classroom pedagogy in the context of this high-stakes testing environment, my colleagues and I observed classrooms using standard observation protocols and closed-ended items to characterize the discourse patterns in classrooms using the distinction between didactic and interactive instruction. Following each lesson, we examined who asked questions most often (teachers or students) and the types of feedback that teachers gave in response to students’ answers (e.g., responses that evaluated correctness only or those that explored students’ comprehension of the correct answer). We also examined the extent to which the students interacted with each other about the course material and the types of questions the students asked during lessons. Examining these discourse patterns allowed me to characterize the classroom pedagogy as interactive or didactic, to gain some understanding of how these schools compared to patterns observed at the beginning of the high-stakes policy (some four to five years earlier), and to compare exposure to these forms of instruction across schools with different racial compositions of students.

Across the case-study schools, the teachers asked most of the questions during the lessons we observed. In 93 percent of the classrooms, “most” or “all” questions were asked by the teachers (combining columns A and B in Table 3). In only 5 percent of the cases were questions asked equally by teachers and students (column C in Table 3). Lessons were typically dominated by teachers, and interactions were between teachers and individual students.
Table 3. Number of Classroom Lessons in Which Questions Were Asked Primarily by Teachers, Primarily by Students, or by Teachers and Students Equally

<table>
<thead>
<tr>
<th>School</th>
<th>Questions Asked Only by the Teacher</th>
<th>Questions Asked Mostly by the Teacher</th>
<th>Questions Asked by the Teacher and Students Equally</th>
<th>Questions Asked Mostly by the Students</th>
<th>Questions Asked Only by the Students</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>31 (34%)</td>
<td>54 (59%)</td>
<td>5 (5%)</td>
<td>0 (0%)</td>
<td>1 (1%)</td>
<td>91</td>
</tr>
</tbody>
</table>

During a typical lesson, the teacher asked the majority of the questions. A vocabulary exercise that preceded a lesson on a story to be discussed during class progressed as follows. The teacher directed students to take out their list of words from the previous class. She then began to ask them questions:

Teacher: What is the first word? (Many students raised their hands)

Teacher: Yes. (The teacher identified a student to respond)

Student: Summertime.

Teacher: Yes. Summertime. What is the second word? (The students were silent as if unsure) What sound does it make? (The students remained still silent)

Teacher: Shanties. The third word? (Many hands went up)

Teacher: Yes. (Identified a student)

Student: Cattails.

Teacher: Yes, cattails (and spelled it out).

This exercise continued for several minutes following a similar pattern.

In a preobservation interview, this teacher said that she was trying respond to the test-based accountability system in this lesson. She explained that “the board and the Iowa test’s emphasis is on higher-order thinking skills . . . interpreting the story, thinking about the characters, students putting themselves into the story.” As a result of these perceived expectations, she was trying to “teach these skills in different ways.” While she sought students’ interpretations, thinking, and connection with the story, she continued to dominate the conversation when the activity shifted from vocabulary to “literature activities,” as the following field notes demonstrate:

In the next activity, [the teacher] went over the handout on “Literature Activities” with the students. . . . [She] did much of the talking, defining the terms, giving her own examples to illustrate the concepts, and occasionally asking students to give their own examples to illustrate the concepts as well. An excerpt from the lesson helps illustrate this pattern.

Teacher: (Identifying a student) “Can you describe your favorite cartoon?”

Student: (silent)

Teacher: “You can’t tell? Let me tell you about mine.” The teacher begins to describe her cartoon. A few students want to interrupt to tell their own stories, but [the teacher] tells them to let her finish her description first. She then calls on two . . . students to give their descriptions (of favorite cartoons). . . . [The teacher] closed off the activity by saying: “That is what Ms.— means when she says describe.”

Therefore, even as the teacher tried to illicit the students’ interpretations and experiences in response to the accountability policy, she tended to dominate the discussion.

In addition to the types of questions the
teachers asked, we also characterized how the teachers responded to the students' answers (see Table 4). For instance, did the teachers evaluate only the correctness of the students' answers, or did they attempt to "dig deeper" by exploring the students' comprehension of the correct answer? The data suggest that teachers respond to students' answers by evaluating their correctness or explaining or elaborating on the correct answer in 69 percent of classrooms (combining columns 3 and 4 in Table 4). This pattern is exemplified by the case cited earlier in which the students' answers were evaluated only for correctness by the teacher. In their responses to the students, most teachers did not push the students toward a conceptual understanding of the class material or ask additional questions to prompt the students to reflect further on the issues under discussion.

The communication in these schools was primarily one-way interaction between the teachers and the students. Across these schools, the students rarely interacted with each other about the class material (see Table 5). In 78 percent of our classroom observations, "no students interacted" or "only a few students interacted occasionally." Thus, the interaction around the course material was almost exclusively between the teachers and individual students, rather than among the students themselves. The students did not have the opportunity to engage in teamwork and joint problem solving and rarely had the opportunity to interact with their classmates around substantive issues. For example, in a representative second-grade classroom, the students read from their journals. During the reading, the teacher reminded them of "their responsibility" to "listen quietly."

In some classrooms, the students were encouraged to interact with each other. One fifth-grade teacher explained that she asked students to agree or disagree with their classmates and to provide explanations for their answers to get students "out of these one-word answers. Give me proof . . . why is this the right answer?" However, this type of instruction was more the exception than the rule in these schools.

While the students were often engaged actively in academic work, they rarely asked substantive questions. Instead, when they asked questions, the questions were most often "procedural" in nature (e.g., asking for clarification of the teacher's instructions), rather than questions in which they sought to clarify and comprehend the "correct" answer or to "gain conceptual understanding" of the instructional material (see Table 6). One example was when the students in a second-grade classroom were given their vocabulary list for the next lesson:

Table 4. Number of Classroom Lessons in Which Teachers Provided Different Types of Feedback to Students' Academic Responses

<table>
<thead>
<tr>
<th>School</th>
<th>No Feedback Provided</th>
<th>Responses Evaluated Only for Correctness</th>
<th>Teacher Explains or Elaborates on Correct Response</th>
<th>Teacher Explains or Elaborates on Knowledge and Comprehension of Correct Response</th>
<th>Teacher Asks Students Another Question to Provide Clues/Prompt Further Thinking</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>2</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Total</td>
<td>3 (3%)</td>
<td>26 (28.5%)</td>
<td>37 (40.6%)</td>
<td>15 (16.4%)</td>
<td>10 (10.9%)</td>
<td>91</td>
</tr>
</tbody>
</table>
Table 5. Number of Classroom Lessons in Which Students Engaged in Various Types of Interactions with Classmates

<table>
<thead>
<tr>
<th>School</th>
<th>No Students Interacted Around the Subject Matter</th>
<th>A Few Students Interacted Occasionally</th>
<th>A Few Students Interacted Frequently</th>
<th>About Half the Students Interacted About Half the Time</th>
<th>Most Students Interacted Occasionally</th>
<th>Most Students Interacted Frequently</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>26 (28.8%)</td>
<td>45 (50%)</td>
<td>3 (3.3%)</td>
<td>8 (8.8%)</td>
<td>5 (5.5%)</td>
<td>4 (4.4%)</td>
<td>90</td>
</tr>
</tbody>
</table>

[The teacher] then began to write a list of words that she asked the students to write on the new sheet of paper. This was their “Vocabulary List” for the new story they would be reading on Monday. As the students asked procedural questions, such as what do with the words, whether they should skip a line in between words, and where to put the title of the story on their papers, she continued to write and asked other students to help whoever was asking the question (by providing the answer). For example, she would say, “Can someone take responsibility to answer that question?” “What did I say, class?”

In this case, the students were focused on understanding what the teacher expected of them, rather than on understanding the course content more deeply. Therefore, across the classrooms, didactic pedagogy was more common than interactive pedagogy.

Does the racial and social class composition of the school matter? I also examined the data by combining schools on the basis of their racial composition to determine whether pedagogy differed by the student composition of classrooms. The schools were divided along racial lines as follows: majority African American (four 100 percent African American schools), majority Latino/Latina (two schools with more than 97 percent Mexican American students), and schools that were racially mixed (more than 40 percent white and combinations of African American, Asian, and Latino/Latina students). For social class, schools were divided into two categories—schools with 90 percent or more students receiving free and reduced-price lunches and schools with fewer than 90 percent of their students receiving free and reduced-priced lunches.

With regard to the kind of feedback that the teachers provided to the students, the teachers in majority African American schools provided the least substantive feedback, while those in mixed-race and majority Latino/Latina schools provided more (see Table 7). More specifically, students in majority Latino/Latina and racially mixed schools were more than twice as likely to have teachers elaborate on correct answers, explore students' knowledge and comprehension of class material (combining columns 5 and 6 in Table 7), and ask follow-up questions to prompt further thinking. This pattern was similar for social class. Teachers in the lowest-income schools provided less substantive feedback to the students.

With regard to students asking questions, those in majority African American schools were the least likely to ask questions, those in majority Latino/Latina schools were the most likely, and those in mixed-race schools were in between. With regard to social class, students in schools with larger percentages of low-income students were more likely to ask questions than those in schools with fewer low-income students (see Table 8).
Table 6. Number of Classroom Lessons in Which Students Asked Various Types of Questions

<table>
<thead>
<tr>
<th>School</th>
<th>Students Rarely Asked Questions</th>
<th>Questions Were Mostly Irrelevant to the Academic Purpose of This Lesson</th>
<th>Questions Were Seeking to Clarify and Comprehend the “Correct” Answer</th>
<th>Questions Were Indicated the Student’s Desire for Conceptual Understanding</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>48 (53.3%)</td>
<td>3 (3.3%)</td>
<td>21 (27.7%)</td>
<td>4 (4.4%)</td>
<td>90</td>
</tr>
</tbody>
</table>

Table 7. Teachers’ Feedback to Students’ Academic Responses, by School Racial Composition

<table>
<thead>
<tr>
<th>School Racial Composition</th>
<th>No Feedback Provided</th>
<th>Responses Evaluated Only for Correctness</th>
<th>Teacher Explains or Elaborates on Correct Response</th>
<th>Teacher Explores Students’ Knowledge and Comprehension of Correct Response</th>
<th>Teacher Asks Another Question to Provide Clues/Prompt Further Thinking</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predominately black</td>
<td>1 (2.5%)</td>
<td>15 (38.4%)</td>
<td>17 (43.5%)</td>
<td>5 (12.8%)</td>
<td>1 (2.5%)</td>
<td>39</td>
</tr>
<tr>
<td>Predominately Latino/Latina</td>
<td>1 (3%)</td>
<td>8 (24.2%)</td>
<td>12 (36.3%)</td>
<td>7 (21.2%)</td>
<td>5 (15.1%)</td>
<td>33</td>
</tr>
<tr>
<td>Predominately mixed</td>
<td>1 (5.2%)</td>
<td>3 (15.7%)</td>
<td>8 (42.1%)</td>
<td>3 (15.7%)</td>
<td>4 (21%)</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>26</td>
<td>37</td>
<td>15</td>
<td>10</td>
<td>91</td>
</tr>
</tbody>
</table>

Table 8. Percentage of Classrooms in Which Students Rarely Asked Questions, by School Racial Composition

<table>
<thead>
<tr>
<th>School Racial Composition</th>
<th>Students Rarely Asked Questions</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predominantly black</td>
<td>25 (65%)</td>
<td>38</td>
</tr>
<tr>
<td>Predominantly Latino/Latina</td>
<td>13 (39.3%)</td>
<td>33</td>
</tr>
<tr>
<td>Racially mixed</td>
<td>10 (52.6%)</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>48 (53.3%)</td>
<td>90</td>
</tr>
</tbody>
</table>
Therefore, African American students may find themselves in the most didactic environments both because they (as students) ask the fewest questions and because their teachers provide them with limited feedback when they respond to the teachers’ questions. Students in overwhelmingly low-income schools are also likely to receive only limited feedback on their responses to teachers’ questions compared to students in schools with smaller percentages of low-income students. Finally, students in predominantly low-income schools ask more questions, but, as was discussed previously, most of these questions are procedural in nature.

These data suggest that in these Chicago classrooms, at least in terms of discourse patterns, teachers ask most of the questions, emphasize correct answers, and elaborate only on correct answers when students provide them. Students focus on understanding what they are being asked to do (rather than on the substantive meaning of the class material) and rarely interact with their classmates around the instructional content during class. Although the Chicago accountability policy is designed to create more interactive environments, data from these classrooms do not suggest that such policies are achieving their goals, particularly for low-income and African American students.

**DISCUSSION AND CONCLUSION**

Some previous research on high-stakes accountability policies has suggested that the policies exert a major influence on instructional content and pedagogy. In response to prior research, I examined teachers’ instructional responses to these policies in Chicago, a district that adopted a high-stakes accountability policy in the mid-1990s. I have argued that teachers’ in the schools and classrooms I studied do not believe that these policies were an overwhelming influence on their instructional practices. Drawing on previous research on the link among the policy environment, school organizations, and classroom instruction, I have shown that high-stakes testing policies are one among many influences on instruction in these schools; that they seem to influence instructional content more often than pedagogy; and that, given the stagnancy of instructional pedagogy in the context of this high-stakes testing reform, racial inequality may not be significantly affected.

**Testing, School Organization, and Instruction**

This article has added to and extended prior work on the link among schools’ external environments, internal organization, and classroom instruction. While neo-institutional theory suggests a decoupling of classroom instruction from the policy environment, the research reported here (in line with some recent neo-institutional work) has shown that the policy environment does shape instruction. However, as prior research suggested, teachers mediate the impact of policies as they make instructional decisions such that preexisting practice often precludes the direct penetration of policy into classroom practice. This article has built on that work by pointing to three additional factors that are at play.

First, influences on instruction are distributed across a number of actors (i.e., other teachers and administrators) and artifacts (i.e., textbooks and the Internet). Thus, the link between the policy environment and instruction is mediated by teachers’ beliefs, experiences, and prior practice, as well as by the constellation of other actors and artifacts within school organizations. Hence, the influence of Chicago’s accountability policy was filtered through teachers’ interactions with teaching colleagues and school leaders and their own beliefs about instruction. This individual and organizational mediation of messages on accountability policy shaped their instruction.

Second, the link between the policy environment and instructional practice varies by the dimension of instruction being examined (in this case, content versus pedagogy). My work has added additional empirical support for the argument that the connection between the policy environment and the classroom depends on the dimension of instruction (Spillane 2005; Spillane and Burch 2006). The teachers in this study were more
likely to report that testing and standards influence the content they teach in comparison to their pedagogy. While previous research has highlighted transformations in pedagogy away from interactive instruction in high-stakes environments (McNeil 2001; Valenzuela 2004), these changes are coupled with state policies that encourage didactic instruction. Therefore, it could be that the nature of the testing regime (which varies by state) can have an impact on the extent to which pedagogy is affected. Moreover, my findings may apply more directly to Chicago during the period when I conducted my research than in the more recent period. When my research was conducted, Chicago used the ITBS (a test not closely aligned with the state standards) to assess students’ progress. More recently, Chicago has moved toward a test (the ISAT) that is more closely aligned to the standards and assesses students’ ability to explain how they arrived at answers in mathematics, for example. This change could have resulted in different influences on pedagogy in the period after my research was conducted.

Third, transforming pedagogy from didactic to interactive runs counter to conventional teaching methods in Chicago classrooms and more generally (Gamoran et al. 2000:57; Goodlad 1984) and requires significant skill. This article has suggested that even when teachers are willing to change how they teach, they may find it difficult to do so. In the Chicago context, most teachers are accustomed to didactic pedagogy, which survey data (Smith et al. 2001) and my classroom observations have demonstrated. Therefore, they are being asked to make major changes in their pedagogy that may prove challenging.

Teachers also seek advice from their teaching colleagues in making these pedagogical changes. However, as my data indicate, most of their teaching colleagues are also primarily didactic instructors. Therefore, the communities of practice in which the teachers are embedded likely provide little guidance in moving from didactic to interactive pedagogy. In fact, they may work against changes in this direction. Therefore, in addition to teachers’ beliefs about teaching and learning and their capacity to make the changes advocated for by the Illinois standards, organizational mediation of policy messages may also limit changes toward more interactive pedagogy. In contrast, moving from interactive to didactic instruction in such contexts may be less challenging.

My work is not meant to be generalized beyond the Illinois and Chicago contexts. For instance, other work has shown that policies that push toward more conventional instruction, characterized by lecture, seat work, and recitation, lead to major changes in pedagogy away from interactive and toward didactic instruction (McNeil 2001; Valenzuela 2004). Other work has suggested that the limited time available for instruction in subjects other than mathematics and language arts (e.g., social studies) in high-stakes testing environments can lead teachers to emphasize more didactic forms of instruction even when they prefer to teach in more interactive ways (Wills 2006). Finally, such policies may play out differently in suburban schools or schools with more affluent student populations because pedagogy in such contexts may have traditionally been more interactive (Anany 1980, 1981). The interplay between the nature of the accountability policy and the individual and organizational mediation processes may account for these different findings.

**Pedagogical Stagnancy and Educational Inequality**

Addressing issues of inequality in education requires focusing on the nature of instruction in urban classrooms. Although other issues also shape inequality, low-income and African American students often receive less demanding instruction than do middle-class and white students. While some standards-based reforms claim to address this issue by transforming what and how students are taught, the teachers in this study reported that these policies do little to transform pedagogy. In fact, previous research has shown that the rhetoric of policies that push for more ambitious instruction often far exceed changes in actual practice (Spillane 2001). Teachers in this study reported that they paid attention to
standards and testing, but that these external policy mechanisms affected the content they covered much more than the way they covered it. My classroom observations demonstrated that students are being exposed to conventional teaching, which is defined as “teaching organized through a set pattern of lecture, recitation, and seat work” (Gamoran et al. 2000:57; see also Goodlad 1984). This kind of instruction is dominated by teachers and geared toward two-way interactions between teachers and individual students even five years after the introduction of Chicago’s high-stakes testing policy. Likewise, the data demonstrate that students in mostly low-income schools and students in predominantly African American schools have the most didactic classrooms. This finding is consistent with data from a survey conducted in the same district during the 1996–97 school year (Smith et al. 2001) and with previous research on the nature of instruction across different types of schools (Ayon 1980, 1981). Lower-income and African American students are not being exposed to valued forms of knowledge that will give them access to highly coveted social positions as adults. Thus standards-based reforms may have little impact on the core issues of pedagogy that many analysts believe will reduce inequality.

**Policy Implications**

On the basis of the findings reported here, there is little evidence that Chicago’s high-stakes testing policy, which targeted both instructional content and pedagogy, had a powerful influence on pedagogy in the classrooms that I studied. What has been consistently demonstrated across multiple studies is that teachers reallocate instructional time to specific content that is likely to be tested and spend more time on preparing for tests. These types of instructional responses are likely to lead to the inflation of test scores, a process through which improvements in students’ test scores outpace the students’ actual learning gains (Koretz 2005). One problem with high-stakes testing policies (as currently implemented) is that while they get teachers’ attention, they provide few resources for addressing issues of inequality in schools. Often, the external supports that are provided are limited in their utility. For instance, while Chicago provided external support for schools on academic probation, this support was judged largely ineffective by objective reviewers (Finnigan, O’Day, and Wakelyn 2003). These policies may also divert public attention away from structural and institutional inequalities that are at the root of many of our contemporary educational challenges (Lipman 2004).

My research suggests that accountability in the system must be coupled with systematic supports for teachers to improve their practice (Darling-Hammond 2004). Simply raising stakes and assuming that teachers’ motivation will lead to needed changes underplays the complexity of transforming instruction. More work is needed to help researchers and educators understand the conditions under which accountability policies affect instruction, which dimension of instruction they influence, and how teachers can be provided with resources to build their capacity to make changes. We also need to gain a better understanding of why instruction varies by the race and social-class composition of schools in ways that reinforce inequality.

In this article, I have argued that the teachers in the Chicago schools I studied reported that high-stakes testing policies have had a limited impact on their pedagogy. In contrast, the teachers reported that they changed the content they covered in response to these policies more frequently. On the basis of these findings, I argue that some advocates and opponents of these policies may have overplayed the policies’ direct impact on instruction. Classrooms are rarely transformed in substantial ways by changes in educational policy. Instead, policies must be made sense of and implemented in local contexts (Spillane 2005; Tyack and Cuban 1995) in which they are mediated by teachers’ existing beliefs and practices and filtered through multiple organizational influences. Perhaps most important, changes are unlikely to result when demands on the system are raised but the resources provided to meet these demands remain stagnant and unequal.
NOTES

1. This section borrows from Diamond and Spillane (2004).
2. Throughout this article, I use ITBS to refer to the Iowa Test of Basic Skills. However, many respondents used the term Iowa to refer to this test.
3. During the period under study, the ITBS was the test used by the Chicago Public Schools to measure school performance. Since 1999, the state of Illinois has used the Illinois Standards Assessment Test (ISAT) to measure school performance. In 2002, the Chicago Public Schools began to use the ISAT and the ITBS to measure school performance. The ITBS and ISAT are thought to be different in their orientation. The ITBS is a norm-referenced test that is designed to measure students’ performance in comparison to a national sample of students. The ISAT is a criterion-referenced test that is designed to measure students’ mastery of the Illinois Learning Standards and is therefore more closely aligned with the Illinois standards than is the ITBS. This shift toward the ISAT may have changed the accountability policy environment in Chicago in the period after my research was conducted.

Analysts suggest that the ISAT is perhaps more challenging than the ITBS because of its longer reading passages, extended response questions, lack of computational problems, and “fewer single-step word problems” in mathematics (Easton et al. 2003:7). However, Easton et al.’s analysis, conducted by the Consortium on Chicago School Research, demonstrated high correlations on students’ performance on these two tests. Given the comparison between these tests, the ISAT could be seen as pushing toward more interactive forms of instruction than the ITBS. As I discuss the findings, I am cognizant that the impact of a test like the ITBS may be different from that of a test like the ISAT. While this is outside the scope of this article, the ISAT could encourage more interactive pedagogy by teachers in comparison to the ITBS. See Easton et al. (2003) for a detailed discussion of the differences between the ITBS and the ISAT.
4. A portion of this comment was also quoted in Spillane, Diamond, and Jita 2003.
5. While classic conceptions of teachers’ work suggest that it is done largely in isolation and with limited collegial interaction (Lortie 1975), the data reported here suggest a higher degree of interaction in these schools.
6. These data are also reported in Spillane et al. (2003).
7. I am not seeking to make causal arguments in discussing these patterns.
8. While the schools’ demographics are reported in the Methods section, I include the categorization scheme used for this particular component of the analysis here to help the reader follow the discussion of my findings.

APPENDIX

Relevant Components of the Classroom Observation Protocol

Cognitive Complexity
(Similar subject matter-specific questions were completed for mathematics and science observations.)
2. The substance of this reading lesson focused mostly on which of the following:
   a. getting students to recall/remember details
   b. getting students to make predictions and/or inferences based on what they read
   c. getting students to articulate their ideas/opinions about the material they read/were about to read
   d. getting students to defend or justify their interpretations
   e. getting students to critique and respond to each others interpretations of text.
   f. getting students to state their response/feelings to the text (e.g., favorite parts).
   g. getting students to apply what they know/have read
APPENDIX

Continued

Classroom Discourse
3. In this classroom during reading lessons, most questions were asked
   a. only by the teacher
   b. mostly by the teacher
   c. by both teacher and students equally
   d. mostly by the students
   e. only by the students

4. What types of questions did the students typically ask during this lesson?
   a. The students rarely asked any questions.
   b. The questions asked were irrelevant to the academic purpose of this lesson.
   c. The questions asked were mostly procedural, such as to clarify instructions.
   d. The questions asked were seeking to clarify and comprehend the “correct” answer.
   d. The questions asked indicated the students’ desire for conceptual understanding.

5. How would you characterize the teacher’s feedback to students’ academic responses?
   a. No feedback provided.
   b. Responses evaluated only for correctness.
   c. The teacher explained or elaborated on the correct response.
   d. The teacher explored the students’ knowledge and comprehension of correct response.
   e. The teacher asked the students another question to provide clues or to prompt further thinking.

6. What types of questions did the teacher ask during the lesson?
   a. The teacher asked no questions or asks only rhetorical questions.
   b. The teacher asked mostly memory or fact questions.
   c. The teacher asked a mix of fact/memory and comprehension questions.
   d. The teacher asked mostly questions that require comprehension or an opinion.

7. Did the questions posed by the teacher vary depending on the particular groups?
   a. Yes, the questions did vary depending on the particular reading group.
   b. No, the questions did not vary depending on the particular reading group.

8. To what extent did the students interact with each other about the subject matter?
   a. No students interacted around the subject matter.
   b. A few students interacted occasionally.
   c. A few students interacted frequently.
   d. About half the students interacted about half the time.
   e. Most students interacted occasionally.
   f. Most students interacted frequently.

Relevant Interview Questions

Questions for Observed Teachers

D1. I noticed that [plug in relevant details from lesson; for example, you immediately informed this student whether her (or his) answer was right or wrong].
   a. Is this something that happens regularly in your [mathematics, science, or reading] teaching?
   b. Why do you do this? Could you do this in another way if you wanted to? If not, why not? [listen for things and/or people that constrain/limit teachers’ choices and control]
   c. Have you always [plug in relevant detail from lesson] in your [mathematics, science, or reading] teaching? If not, how long have you done it this way? What did you do
APPENDIX
Continued

in the past? How did you change? Why? Did anyone or anything contribute to this change? Who? How and what did they contribute? Why this person? or Why this thing?

Questions for Nonobserved Teachers
D1. Are there particular things about the way you teach mathematics/science/literacy now that you are unhappy with? If yes, what are these things? What makes you unhappy about [list aspect of mathematics instruction identified by the informant]? [If no, skip to question D4]
D2. How long have been unhappy about [list aspect of mathematics/science/literacy instruction identified by the informant]?
D3. Was there something or someone that helped you see your teaching of mathematics/science/literacy or [list aspect of mathematics instruction identified by the informant] as problematic? Who? What? How?
D4. Have you changed anything about the way you teach mathematics/science/literacy recently? [Try to focus the respondent on content, materials, teaching strategies, and grouping arrangements as specific aspects of instruction]. [If yes, continue with D5. If no, go to D8]
D5. If you can remember back to when you made this change, was there someone or something that helped you to think about adopting this new way of doing things? Who? How?
D6. Did anyone help you make this change in your classroom? Who? How did this person contribute to them? Why do or did you turn to this person?
D7. Did you draw on any resources to help you make this change? What resources did you draw on?

REFERENCES


Gamoran, Adam, Walter G. Secada, and Cora B.


John B. Diamond, Ph.D., is Assistant Professor of Education, Graduate School of Education, Harvard University. His main fields of interest are education, racial and class stratification, leadership and organizational change, and policy implementation. He is currently studying race, class, and school achievement in multiracial suburbs. He is coeditor of Distributed Leadership in Practice (Teachers College Press, 2007)

Work on this paper was supported by the American Educational Research Association/Institute for Education Sciences Research Grant Program, the Harvard University William F. Milton Fund, the Distributed Leadership Project funded by research grants from the National Science Foundation (REC-9873583) and the Spencer Foundation (200000039), the National Academy of Education/Spencer Postdoctoral Fellowship Program, and the Radcliffe Institute for Advanced Study. The Harvard Graduate School of Education also supported this work. The author thanks Kristy Cooper, Shannon Hodge, Loyiso Jita, Brenda Lin, Amy Coldren, Fred Brown, Patricia Burch, Tim


Hallett, Curtis Askew, Adam Roux, and Ryon Lancaster for assistance with data collection, management, and analysis and Walter Allen, Betty Malen and Heinrich Mintrop for their helpful comments on earlier drafts. Previous versions of this article were presented at the annual meeting of the American Educational Research Association, Montreal, Canada, 2005; the annual meeting of the Association of Black Sociologists, Philadelphia, 2005; and the annual meeting of the American Sociological Association, Montreal, Canada, 2006. All opinions and conclusions expressed in this article are those of the author and do not necessarily reflect the views of any funding agency or institution. Address correspondence to John B. Diamond, Harvard Graduate School of Education, 6 Appian Way, 418 Gutman Library, Cambridge, MA 02138; e-mail: diamonjo@gse.harvard.edu.