How do principals adjust to new standards and measurement rulers of school performance when they are coupled with high-stakes consequences? Adopting a contingent view of leadership and a multiple accountabilities framework, this qualitative study explores how high school principals enacted their accountability in New York City, by comparing two periods: just before such policy changes took effect and just after. We observed large differences in principals' attention to external political, bureaucratic, and market accountabilities; how much accountability conflict or alignment was experienced; and how principals' sense of professionalism was constructed. These differences may portend a new politics of the principalship.

Principals have a power advantage over other school actors. Betty Malen and Melissa Cochran (2008) argued that they typically use their influence to protect established interests and deflect criticism by “avoiding, suppressing and containing conflict” (p. 168). Yet recent policy shifts may be narrowing the scope of principal influence or providing opportunities for principals to orchestrate conflict rather than suppress it.

One arena of policy change is accountability. Principal success or failure is increasingly evaluated from the perspective of reformers who aim to alter their behavior as a means of achieving targeted educational improvements (Wong & Nicotera, 2007). Nowhere is this more evident than in New York City. Each high school principal's performance rating depends upon the school's yearly grade, calculated as a function of its attendance, testing, and graduation ranking when compared to "peer" schools. The school grade either adds points to the principal's performance rating or starts her at a deficit. Poorly rated principals can be removed in as little as two years; highly rated ones can earn significant bonuses (New York City Department of Education, 2008a).
Despite the strong incentives in such evaluation systems, their operation is neither mechanistic nor uncomplicated (Bizar & Barr, 2001; Mulford, 2008; O'Day, 2002). The politics of school accountability are so complex and uncertain that tightly linked systems of behavior, rewards, and punishment may serve to relegate the public schools even if they do not change the way leaders think and act (Malen & King-Rice, 2009). To explore how accountability affects leaders' behavior, we examine these politics from the inside out, that is, from the principal's perspective. We ask, How do New York City high school principals construct their accountability environments? How do the types of schools they lead influence their experience of accountability? What changes in the perceived accountability environment when consequences are stiffened? How might this alter the politics of the principalship?

To study these interrelated issues we interviewed two samples of high school principals in New York City, 3 years apart. We chose high schools because recent research suggests that their sense of efficacy operates differently from that of elementary school principals’ (Leithwood & Jantzi, 2008, p. 524), teacher cultures differ greatly between elementary and high schools (Walstrom & Louis, 2008), and because high school principals perceive significantly more job-related problems than do elementary or middle school principals (Gates, Ringel, Santibanez, Ross, & Chung, 2003; Teddlie, Stringfield, & Reynolds, 2000). Moreover, in New York City they confront performance targets quite different from those designed for elementary school principals.

Our first interviews took place in 2004-05, 2 years after mayor Michael Bloomberg took control of the schools in 2002, closed 36 subdistrict offices, and abolished all school boards. Bloomberg also mandated a common curriculum for reading and math; reorganized the operational bureaucracy; and created 10 new instructional units, called regions, made up of 100 to 150 schools each (Fruchter, 2008; Viteritti, 2009). Under these centralizing governance reforms, principals were exhorted to take responsibility for the performance of their schools, but the city's performance-tracking mechanisms were still in development. State Report Cards—relying on test scores from the Regent’s Examinations that all New York State students must pass to graduate—were the primary means of assessing school performance.¹ There were no special consequences for principals unless a school was judged so poor it must be shut down by the state.

The second round of interviews took place in the school year 2007-08, just after the Bloomberg administration had implemented three new school performance measures and tied them to principals’ performance ratings. After 2006, the city mandated three district-developed measures of school performance. Each high school’s Progress Report (PR) weighs several types of performance data: Regent’s exams (30%), the progress of the high school in “moving” students to higher levels of performance and on-time graduation (55%), and attendance combined with the results of surveys (15%).² Learning Environment Surveys (LES) assess the satisfaction levels of three constituencies—parents, teachers, and students—on school safety and respect, academic expectations, engagement, and communication. PRs also identify whether the school received “extra credit” for unpredicted performance gains among English language learners, special education, and/or its lowest scoring students. Beginning in 2006-07 these measures were summed into a single score and transformed into a school grade, A–F. School Quality Reviews (SQRs)

¹As of 2007-08, to receive a high school diploma all students were required to pass five Regent’s exams, one each in English, math, science, global studies, and American history.
²As of 2008-09, these relative weights shifted to 25%, 60%, and 15%, respectively.
resulted from 2-day inspection visits conducted yearly by contracted educational consultants. SQR inspectors graded schools as “Undeveloped,” “Proficient,” or “Well-Developed.”

Since 2007, principals have been graded on a newly revamped Professional Personnel Review that relies on the PR, the SQR, the Learning Environment Surveys, and up to five goals set by the principal. Principals are told they have 2 years to meet district-set performance goals or risk losing their jobs. Principals are also eligible to receive bonuses of up to $25,000 if their PR scores are in the top 20% of the city. In fall of 2008, three high school principals received this top prize, and another 30 received a lesser bonus (Hernandez, 2008b).

We did not anticipate such dramatic changes when the first interviews were undertaken. By 2007, however, we realized our common research protocol made it possible to reanalyze our data for insights into principals’ actions in the face of extreme shifts in accountability policy. Notwithstanding our inability to conduct formal pre-postanalysis because of differences in our two sampling frames, our findings are strongly suggestive. If confirmed, they may portend a new politics of the principalship.

FRAMING ACCOUNTABILITY AND LEADERSHIP

Accountability consists of demands—from oneself or legitimate others—for a demonstration of performance claims. For example, politicians manifest accountability by listening to and addressing the needs of their constituents. Managerial employees show their bosses that they have met established targets, and are using preferred means to do so. Market-based entrepreneurs please their customers, or risk losing them to another provider. Lawyers demonstrate that they follow professionally accepted practice or risk ostracism and censure, while religious leaders live in conformance with the teachings of their faith.

School leaders face all of these demands. Principals show they listen to and address the problems of parents and community members by negotiating conflict or by mobilizing support in anticipation of their demands (Malen & Cochran, 2008). Principals manage the gaps between external standards of school performance and current performance—test targets, but also procedural requirements—by managing resources and coordinating expectations so that the standards become common goals (Carnoy, Elmore, & Siskin, 2003). In response to market pressures, principals compete among themselves for excellent students, teachers, and other scarce resources. When asked to demonstrate their professionalism in curricular and instructional leadership, they are called upon to coach and mentor their subordinates (Stein & Nelson, 2003). Like religious leaders, school principals use moral persuasion and ethical reasoning to justify their decisions (Sergiovanni, 1992). All these accountability claims have been previously identified in typologies (Adams & Kirst, 1999; Darling-Hammond, 1989), and by factor analysis that identifies multiple components of successful principal decision making (Leithwood & Jantzi, 2008; Mulford, 2008).

Table 1, adapted from Firestone and Shipps (2005), outlines how each of five accountability types might be expected to influence principal behavior.

---

3In 2007–08, SQRs were graded as “outstanding,” “well-developed,” “proficient,” “undeveloped with proficient features,” or “underdeveloped.”

4Schools must volunteer to be in the bonus eligible pool, but all principals are subject to the grading sanctions.
A NEW POLITICS OF THE PRINCIPALSHIP? 353

TABLE 1
Typology of School Leaders' Accountabilities

<table>
<thead>
<tr>
<th>Type</th>
<th>Manifestations</th>
<th>Objectives</th>
<th>Operative Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Political</td>
<td>Citizen pressure</td>
<td>Satisfaction</td>
<td>Suppressing, negotiating or mobilizing conflict</td>
</tr>
<tr>
<td>Bureaucratic</td>
<td>Process</td>
<td>Regulations</td>
<td>Compliance</td>
</tr>
<tr>
<td></td>
<td>Outcome</td>
<td>Standards and consequences</td>
<td>Managing resources and coordinating expectations</td>
</tr>
<tr>
<td>Market</td>
<td>Entrepreneurialism</td>
<td>Creativity, efficiency</td>
<td>Competition</td>
</tr>
<tr>
<td>Professional</td>
<td>Practical consensus</td>
<td>Preferred practice</td>
<td>Coaching and mentoring</td>
</tr>
<tr>
<td>Moral</td>
<td>Espoused beliefs</td>
<td>Value commitments</td>
<td>Moral persuasion or ethical reasoning</td>
</tr>
</tbody>
</table>

Note. Adapted from Firestone and Shipps (2005).

As Table 1 indicates, reform efforts create new, overlapping, and sometimes competing accountability obligations. For example, principals are expected to achieve bureaucratically targeted goals first, even if this means putting off other educational objectives that teachers, parents, and students value (see, e.g., Linn, Baker, & Betebenner, 2002; Sleeter, 2007). Yet, in New York City as elsewhere, open enrollment and the creation of hundreds of small high schools has pitted principals against one another in competition for students whose prior performance predicts they can meet outcome expectations (New York City Department of Education, 2008b). At the same time, parents are clamoring for more voice in school disciplinary policies, criticizing school leaders for a lack of consultation (Commission on School Governance, 2008). Accrediting associations, teacher-training institutions, and teachers also do not easily forgive what they perceive to be nonstandard professional practice even when it is initiated in response to outcome goals. Fairness remains important; many urban school systems, including New York City's, serve a highly diverse student body and are located in economically stratified cities.

School principals' accountability tasks are also situation specific. One recent study based on a multiple accountabilities framework concluded that variations in influence by state agencies, parents' associations, teachers, and school councils on instruction and supervision decisions significantly impacted how principals interpreted their own power and influence. Teachers and parents' associations seemed to pull principals in opposite directions; strong teacher influence diminished principals' willingness (or ability) to be influenced by parent groups and vice versa (Marks & Nance, 2007, p. 25).

Contingent and goal-directed leadership theories better capture our view of the principalship than do trait, skill, or style theories (Northouse, 2006). We adopt a contingent approach in which leaders perceive their responsibilities as varied depending upon the relationships they have with adults crucial to their organization's performance, the structure of the educational task, and the positional power they have to reward or coerce desired behavior. At the same time, we assume principals are goal oriented; they must motivate teachers, students, and parents to achieve outcomes that are frequently challenging.

In this view, principals do not have the luxury of simply responding to the loudest or even the most immediate accountability pressure. They also do not simply apply their acquired skills
or leadership styles, as much of the survey research assumes. Instead, principals construct their accountability environments, focusing on some concerns over others, balancing competing demands, and inevitably making choices (Firestone & Shipps, 2005). One principal may experience zero-sum choices between competing sets of accountability demands, even constructing her accountability environment so that two are always in existential conflict. Another principal, or the same principal at a different time, may experience the same accountability demands as aligned: several types of accountability reinforcing one another. Still others may focus on a single accountability type that has grown distinctive enough in particular school environments to crowd out all others (White, 2006).

Before we analyzed our data, we developed hunches about the differences we anticipated with respect to the school’s adult relationships, the structure of the educational task, and the leader’s positional authority. We measured variations in adult relationships by the school size and the principal’s gender. We measured variations in the school’s educational task by the level of student diversity and by students’ aggregate level of pre-high school achievement. These two dimensions of leadership are assumed to be stable factors of influence, changing little, if at all, over the 3 years between 2004-05 and 2007-08.

We measured variations in the principal’s positional power by contrasting the ways they enacted their accountability environments in 2004-05 and 2007-08. This measure would directly respond to the shifts in district policy that the New York City Department of Education (DOE) adopted but also serve as a proxy for principals’ perceived scope of influence. We attended to shifts in principals’ descriptions of the types of accountability that dominated their environment, how principals’ views changed about whether accountabilities were in conflict, aligned, or unopposed, and which specific combinations they identified as crucial to their decision making.

We found strong evidence that principals reconstruct their accountability environments when the district changes its performance measures and consequences but were relatively unaffected by the adult relationship and educational task variables we used. These findings may be artifacts of differences in the two samples—alternative hypotheses we note throughout—but we argue that principals own descriptions support a policy-based explanation of the changes we found. If so, our results complicate an emerging argument that external demands of various kinds are dominating principals’ perceptions of their authority largely by reducing their perceived discretion (Malen & Cochran, 2008). We found a dramatic increase in the sheer numbers of external demands that New York City principals attend to, but the amount of conflict between these demands and principals’ sense of professionalism diminished over the same period. New York City principals now appear to acknowledge a wider array of external actors as legitimate stakeholders, and many are beginning to call upon them for assistance in meeting educational goals. If substantiated, our findings suggest a turning point in the politics of the principalship. But, before we examine the details, we describe our studies.

METHODS

The 10 principals sampled in 2004-05 were purposively selected based on their independently judged reputation for leading “good” schools, defined in terms of instructional competence,
using two of Larry Cuban's (2003) categories—progressive and traditional. This resulted in 10 high schools and an initial response rate of 33%. These respondents were supplemented with six additional high schools using snowball methods. The 18 2007–08 principals were selected from a stratified random sample and offered a $50 gift certificate for their school as an incentive to participate. An initial response rate of 50% required that the process of random selection be repeated. New York City has a sizeable and growing number of small high schools. Consequently, both waves were stratified to include equal numbers of small and large high schools.

All interviews followed the same protocol, developed by the authors and a colleague, Tarik Razik, in 2004 (Razik, 2006; White, 2006). The interview protocol was based on the Firestone and Shipps (2005) framework of multiple accountabilities and adapted critical incident technique (Flanagan, 1954). Critical incident technique’s open-ended questioning is not biased toward one or another theoretically driven view of leadership (Northouse, 2006), and it has proven useful for identifying actors’ perceptions of behaviors critical to their own performance in environments where individual discretion is high and decisions have serious consequences (Chell, 1998), like New York City’s high stakes accountability environment.

Each principal was contacted to secure an appointment at the school site. Interviews were tape-recorded, transcribed, and coded for critical incidents. Permission to enter the schools was secured as required by college and DOE institutional review processes. All interviews were conducted confidentially. Names provided herein are pseudonyms and some descriptive data are omitted or changed to preserve anonymity.

Our contingent and goal-directed framing of leadership requires active sense making on the part of school leaders. Consequently, we define an incident as a reflection of the principal’s perspective, not as a reliable statement of what improves performance. Often it is related as a story, with a beginning, middle, and end, and/or discernable causes and predicted consequences. This framing has intentional resonance with the narrative interests of political policy analysts, who judge political behavior as much by the stories that are told to justify, explain, or rationalize it, as by the actions that unambiguously result (Stone, 1998). For our purposes, a critical incident is any situation involving a principal decision that, in hindsight, seemed consequential. Leaders may feel certain that a particular decision on their part was self-evidently purposive, or that change in their working environment obviously shifted the consequences they face—believing

---

5Two additional categories of schools mentioned by Cuban—democratic and community based—were not sufficiently distinguishable to be useful.

6Three “slots” required only one new draw, three required a second redraw, one required three redraws. The last slot required six redraws. After testing the data for differences by race, gender, school performance rating, and several other characteristics to determine if the schools requiring a redraw were systematically different from those responding to the original invitation, we concluded that the most plausible reason for the redraws was simply time. The longer it took to schedule an interview, the less likely the principal was to accept the invitation, since the school term became increasingly busy with spring testing, budgetary reconciliation, and graduation.

7In both protocols, the population excluded schools with examination or audition entrance criteria (n = 7), schools for overage students, GED-only programs, schools assigned to incarcerated students and those labeled by the State of New York as “In Need of Improvement” for 2 years in a row.

8The 2007–8 sample was coded using NVivo software.
this incident was *critical*—even when their intentions and/or the shifted incentives do not reliably have a causal relationship to outcomes.\(^9\)

The interview protocol consisted of seven open-ended questions structured to elicit incidents critical to the principal’s experience of accountability at the time of the interview. Principals were asked to provide an explanatory example, or typical story, about each type of accountability, although there was no limit on the number of incidents they could relate and the range proved wide (\(n = 5\)–17). Multiple probes ensured that critical incidents were described in detail (e.g., Was anyone else involved in this issue? Did you consult anyone about what to do? What was their response? etc.). Interviews lasted, on average, 1.5 hrs.

We coded each critical incident for the accountability represented and coded incidents were the unit of analysis. Reliability was assured in two ways: The protocol was pilot tested and revised by the second author and a colleague in 2004. Interrater reliability was assured in 2004 by an initial agreement between two coders of 64% with a Cohen’s kappa (taking account of chance) of .55, considered “moderate” by Stemler (2001). In 2008 the authors of this study initially achieved 75% agreement in their coding of a sample of incidents and a Cohen’s kappa of .41—again, “moderate” consensus. Our samples are not large enough to meet assumptions of statistical validity, but we expect this study will help us understand how major policy changes affect what principals are able and willing to attend to in their accountability environments, and help us understand when and how principals feel powerful.

### Characteristics of the Schools and Principals

Table 2 summarizes some characteristics and school accountability measures that describe the two samples. The schools were, on average, similar in terms of student race and ethnicity, although the 2004–05 schools were somewhat more highly diverse, whereas the 2007–08 schools evidenced more variation. Unsurprisingly, the peer index, the average attendance rate, the rate of freshman on track for graduation, and the 4-year graduation rate were all somewhat higher in the 2004–05 schools, which were selected based on their “good” reputation. The 2004–05 schools also educated far fewer English Language Learners on average than did the 2007–08 schools.\(^10\) All the schools had the traditional configuration of Grades 9 to 12, although three schools in the 2007–08 group also had Grades 6 to 8.

In one important way the principals differed: In 2004–05 their average tenure was more than twice as long as the average tenure of those interviewed in 2007–08. Our interviews provide a great deal of anecdotal evidence that there are fewer seasoned principals in the system today. In 2007–08, some of our informants told us they intended to leave the principalship because the policy changes had so altered the job. But because the first sample was selected purposively and the second sample randomly, it may also be that the stark differences are a sampling artifact. In any

---

\(^9\) Some studies find low correlations between principal's high self-ratings on these components and school outcomes. Leithwood and Jantz (2008) reported that in path analysis self-efficacy does not have an independent effect on student outcomes.

\(^10\) Other school characteristics reflected this pattern of little or no average difference between samples except when the characteristic was a calculated score reflecting performance or ability to produce higher performance.
<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Sex</th>
<th>Race (not self-reported)</th>
<th>Years as Principal (at time of interview)</th>
<th>% Black</th>
<th>% Hispanic</th>
<th>% White</th>
<th>% Asian/PI</th>
<th>Diversity Index</th>
<th>% Peer Index</th>
<th>Attendance</th>
<th>School Size</th>
<th>Grades</th>
<th>Freshman On Track</th>
<th>City's 4-yr grad rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemlock</td>
<td>F</td>
<td>W</td>
<td>6</td>
<td>33%</td>
<td>59%</td>
<td>5%</td>
<td>3%</td>
<td>0.7</td>
<td>3%</td>
<td>2.53</td>
<td>81%</td>
<td>Small</td>
<td>9-12</td>
<td>65%</td>
</tr>
<tr>
<td>Oak</td>
<td>F</td>
<td>W</td>
<td>10</td>
<td>34%</td>
<td>64%</td>
<td>1%</td>
<td>0%</td>
<td>0.6</td>
<td>10%</td>
<td>2.35</td>
<td>83%</td>
<td>Small</td>
<td>9-12</td>
<td>70%</td>
</tr>
<tr>
<td>Fir</td>
<td>F</td>
<td>B</td>
<td>1</td>
<td>76%</td>
<td>17%</td>
<td>2%</td>
<td>4%</td>
<td>0.5</td>
<td>1%</td>
<td>2.8</td>
<td>88%</td>
<td>Large</td>
<td>9-12</td>
<td>49%</td>
</tr>
<tr>
<td>Birch</td>
<td>F</td>
<td>W</td>
<td>20</td>
<td>9%</td>
<td>58%</td>
<td>13%</td>
<td>20%</td>
<td>0.7</td>
<td>5%</td>
<td>2.97</td>
<td>92%</td>
<td>Large</td>
<td>9-12</td>
<td>75%</td>
</tr>
<tr>
<td>Dogwood</td>
<td>M</td>
<td>W</td>
<td>20</td>
<td>38%</td>
<td>25%</td>
<td>33%</td>
<td>4%</td>
<td>0.9</td>
<td>1%</td>
<td>N/A</td>
<td>N/A</td>
<td>Small</td>
<td>9-12</td>
<td>N/A</td>
</tr>
<tr>
<td>Palm</td>
<td>M</td>
<td>W</td>
<td>20</td>
<td>38%</td>
<td>44%</td>
<td>14%</td>
<td>4%</td>
<td>0.8</td>
<td>2%</td>
<td>N/A</td>
<td>76.3%*</td>
<td>Small</td>
<td>9-12</td>
<td>N/A</td>
</tr>
<tr>
<td>Willow</td>
<td>M</td>
<td>W</td>
<td>3</td>
<td>10%</td>
<td>63%</td>
<td>18%</td>
<td>8%</td>
<td>0.7</td>
<td>2%</td>
<td>2.73</td>
<td>91%</td>
<td>Small</td>
<td>9-12</td>
<td>N/A</td>
</tr>
<tr>
<td>Redwood</td>
<td>M</td>
<td>W</td>
<td>10</td>
<td>19%</td>
<td>24%</td>
<td>51%</td>
<td>6%</td>
<td>0.8</td>
<td>2%</td>
<td>3.44</td>
<td>95%</td>
<td>Large</td>
<td>9-12</td>
<td>96%</td>
</tr>
<tr>
<td>Pine</td>
<td>M</td>
<td>W</td>
<td>1</td>
<td>34%</td>
<td>14%</td>
<td>21%</td>
<td>31%</td>
<td>0.9</td>
<td>2%</td>
<td>3</td>
<td>93%</td>
<td>Large</td>
<td>9-12</td>
<td>85%</td>
</tr>
<tr>
<td>Banyon</td>
<td>M</td>
<td>W</td>
<td>10</td>
<td>50%</td>
<td>22%</td>
<td>3%</td>
<td>25%</td>
<td>0.8</td>
<td>11%</td>
<td>2.62</td>
<td>84%</td>
<td>Large</td>
<td>9-12</td>
<td>60%</td>
</tr>
<tr>
<td>2004-06 Sample Avg/Prop</td>
<td>40% female</td>
<td>10% minority</td>
<td>10.1</td>
<td>34%</td>
<td>39%</td>
<td>16%</td>
<td>11%</td>
<td>0.7</td>
<td>4%</td>
<td>2.81</td>
<td>88%</td>
<td>50% Small</td>
<td>69%</td>
<td></td>
</tr>
<tr>
<td>2004-05 Sample SD</td>
<td>7.26</td>
<td></td>
<td></td>
<td>0.1</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proust</td>
<td>F</td>
<td>W</td>
<td>4</td>
<td>74%</td>
<td>21%</td>
<td>1%</td>
<td>1%</td>
<td>0.5</td>
<td>1%</td>
<td>2.88</td>
<td>92%</td>
<td>Small</td>
<td>9-12</td>
<td>91%</td>
</tr>
<tr>
<td>Ishiguro</td>
<td>F</td>
<td>B</td>
<td>4</td>
<td>42%</td>
<td>52%</td>
<td>3%</td>
<td>2%</td>
<td>0.7</td>
<td>8%</td>
<td>2.37</td>
<td>82%</td>
<td>Small</td>
<td>9-12</td>
<td>80%</td>
</tr>
<tr>
<td>Steinbeck</td>
<td>F</td>
<td>B</td>
<td>4</td>
<td>83%</td>
<td>12%</td>
<td>1%</td>
<td>3%</td>
<td>0.4</td>
<td>1%</td>
<td>2.64</td>
<td>83%</td>
<td>Small</td>
<td>9-12</td>
<td>72%</td>
</tr>
<tr>
<td>Salingar</td>
<td>F</td>
<td>B</td>
<td>4</td>
<td>37%</td>
<td>61%</td>
<td>0%</td>
<td>2%</td>
<td>0.6</td>
<td>7%</td>
<td>2.6</td>
<td>84%</td>
<td>Large</td>
<td>9-12</td>
<td>66%</td>
</tr>
<tr>
<td>Angelos</td>
<td>F</td>
<td>W</td>
<td>4.5</td>
<td>59%</td>
<td>34%</td>
<td>2%</td>
<td>0%</td>
<td>0.6</td>
<td>7%</td>
<td>2.41</td>
<td>71%*</td>
<td>Large</td>
<td>9-12</td>
<td>53%</td>
</tr>
<tr>
<td>Murdock</td>
<td>F</td>
<td>A</td>
<td>6</td>
<td>30%</td>
<td>48%</td>
<td>3%</td>
<td>17%</td>
<td>0.8</td>
<td>13%</td>
<td>2.75</td>
<td>83%</td>
<td>Large</td>
<td>9-12</td>
<td>46%</td>
</tr>
<tr>
<td>Marquez</td>
<td>F</td>
<td>W</td>
<td>5</td>
<td>28%</td>
<td>35%</td>
<td>6%</td>
<td>26%</td>
<td>0.9</td>
<td>10%</td>
<td>2.52</td>
<td>75%</td>
<td>Small</td>
<td>9-12</td>
<td>N/A</td>
</tr>
<tr>
<td>Shelley</td>
<td>M</td>
<td>W</td>
<td>1</td>
<td>45%</td>
<td>54%</td>
<td>1%</td>
<td>0%</td>
<td>1%</td>
<td>3%</td>
<td>N/A</td>
<td>N/A</td>
<td>Small</td>
<td>9-12</td>
<td>N/A</td>
</tr>
<tr>
<td>Tan</td>
<td>M</td>
<td>W</td>
<td>5</td>
<td>14%</td>
<td>45%</td>
<td>16%</td>
<td>24%</td>
<td>0.8</td>
<td>67%</td>
<td>2.75</td>
<td>92%</td>
<td>Small</td>
<td>9-12</td>
<td>78%</td>
</tr>
<tr>
<td>Catter</td>
<td>M</td>
<td>W</td>
<td>6</td>
<td>40%</td>
<td>52%</td>
<td>6%</td>
<td>2%</td>
<td>0.7</td>
<td>5%</td>
<td>2.56</td>
<td>89%</td>
<td>Small</td>
<td>9-12</td>
<td>71%</td>
</tr>
<tr>
<td>Hemingway</td>
<td>M</td>
<td>W</td>
<td>3</td>
<td>51%</td>
<td>35%</td>
<td>4%</td>
<td>3%</td>
<td>0.7</td>
<td>3%</td>
<td>2.74</td>
<td>87%</td>
<td>Small</td>
<td>9-12</td>
<td>55%</td>
</tr>
<tr>
<td>Morrison</td>
<td>M</td>
<td>W</td>
<td>6</td>
<td>89%</td>
<td>9%</td>
<td>0%</td>
<td>0%</td>
<td>0.2</td>
<td>1%</td>
<td>2.65</td>
<td>88%</td>
<td>Small</td>
<td>9-12</td>
<td>65%</td>
</tr>
<tr>
<td>Tenayson</td>
<td>M</td>
<td>W</td>
<td>5</td>
<td>35%</td>
<td>57%</td>
<td>1%</td>
<td>6%</td>
<td>0.7</td>
<td>13%</td>
<td>2.48</td>
<td>84%</td>
<td>Small</td>
<td>9-12</td>
<td>33%</td>
</tr>
<tr>
<td>Broste</td>
<td>M</td>
<td>W</td>
<td>6</td>
<td>86%</td>
<td>12%</td>
<td>1%</td>
<td>0%</td>
<td>0.3</td>
<td>1%</td>
<td>2.72</td>
<td>85%</td>
<td>Large</td>
<td>6-12</td>
<td>70%</td>
</tr>
<tr>
<td>Gates</td>
<td>M</td>
<td>W</td>
<td>4</td>
<td>13%</td>
<td>18%</td>
<td>59%</td>
<td>9%</td>
<td>0.7</td>
<td>6%</td>
<td>2.74</td>
<td>83%</td>
<td>Large</td>
<td>6-12</td>
<td>66%</td>
</tr>
<tr>
<td>Swinburne</td>
<td>M</td>
<td>B</td>
<td>1</td>
<td>59%</td>
<td>18%</td>
<td>1%</td>
<td>19%</td>
<td>0.7</td>
<td>14%</td>
<td>2.32</td>
<td>80%</td>
<td>Large</td>
<td>9-12</td>
<td>35%</td>
</tr>
<tr>
<td>Elliot</td>
<td>M</td>
<td>W</td>
<td>5</td>
<td>9%</td>
<td>49%</td>
<td>15%</td>
<td>27%</td>
<td>0.8</td>
<td>23%</td>
<td>2.6</td>
<td>84%</td>
<td>Large</td>
<td>9-12</td>
<td>50%</td>
</tr>
<tr>
<td>Ellison</td>
<td>M</td>
<td>W</td>
<td>6</td>
<td>29%</td>
<td>21%</td>
<td>16%</td>
<td>33%</td>
<td>0.9</td>
<td>14%</td>
<td>2.85</td>
<td>88%</td>
<td>Large</td>
<td>9-12</td>
<td>42%</td>
</tr>
<tr>
<td>2007-08 Sample Avg/Prop</td>
<td>38% female</td>
<td>27% minority</td>
<td>4.42</td>
<td>46%</td>
<td>35%</td>
<td>8%</td>
<td>10%</td>
<td>0.6</td>
<td>11%</td>
<td>2.63</td>
<td>84%</td>
<td>50% Small</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>2007-08 Sample SD</td>
<td>1.54</td>
<td></td>
<td></td>
<td>0.2</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Student Race/ethnic breakdown based on 2005-06 data.  
N/A = data not available because the school was starting up this year and had no previous record with which to make comparisons and/or did not have 4 years of operation required for graduation rates.  
*= data based on average of several rates provided in report card.
case, the lack of overlap precluded our analysis of principals’ responses by years of experience, one alternative measure of adult relationships.

HOW DO PRINCIPALS DESCRIBE CRITICAL INCIDENTS?

The 10 principals interviewed in 2004–05 reported a total of 96 critical incidents, while the eighteen principals in 2007–08 reported 241 critical incidents. As we discuss in our findings, the larger average number of critical incidents in the later years may suggest that stepped up sanctions and incentives by 2007–08 had made more types of incidents seem critical.

Our 2004–05 informants identified 29 different combinations of accountability types and relationships between them, whereas in 2007–08 the range consisted of 31 distinct combinations of relationships of conflict and support. Every principal identified at least one example of each type of accountability—political, bureaucratic, professional, market, and moral—and all identified several instances of conflict between accountabilities that principals described as forcing a choice between two competing pressures. Ninety-three percent of principals experienced accountability incidents that were complex, in which three types of accountabilities were pulling them in different directions or where one kind of accountability supported another so that a principal’s decision making involved at least three constituencies and at least two different decision options. At the other end of the spectrum, 85% reported experiencing accountabilities that were aligned: two or more accountability types supporting the same decision. Only slightly more than half (57%) experienced uncontested types of accountability, with no competing pressures and no supporting ones either. Strikingly, all these incidents come from 2007–08.

What follows is a summary of the top 91% of all incidents. We attenuated the full list because the bottom 9% represent accountability types and relationships that each occurred in less than one percent of incidents. Most are one-off combinations that reflected idiosyncratic coding or one principal’s unique experiences with accountability; the long “tails” of these two distributions.

DISCONFIRMING EVIDENCE AND UNANTICIPATED SURPRISES

Our initial hunches about gender, size, school diversity, and students’ prior performance were all based on more or less commonly accepted distinctions in the research literature. They proved, at best, weakly substantiated by our evidence. We briefly summarize these weak associations before exploring our key findings in greater detail.

Table 3 displays the major findings of two analyses on all 28 principals interviewed. The far left column shows the types of accountability identified (i.e., political, bureaucratic, professional, market, or moral) as well as the relationships between them: Unconflicted types are reported with the full name, simple conflict between two types of accountability is shown by a “vs.” between the two, simple support is represented by a “supp” between two types of accountability, and complex conflict and/or support is represented by three types of accountability with a mixture of “vs.” and/or “supp” between them. The second two columns provide the number of incidents coded by type of accountability, and the proportion of all 306 incidents mentioned. The first two comparisons (columns 4–7) show how principals constructed their accountability environments when sorted by our measures of adult relationships: principal gender and school size. The last two
TABLE 3
How Do Variations in School Relationships and Educational Tasks Affect Principals' Experiences of Accountability?

<table>
<thead>
<tr>
<th>Incident Types</th>
<th>Count</th>
<th>%</th>
<th>Gender</th>
<th>Size</th>
<th>Structure of the Education Task</th>
<th>Peer Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Female</td>
<td>Male</td>
<td>Diversity Index</td>
<td>Bottom half</td>
</tr>
<tr>
<td>Combined Samples</td>
<td>n = 28 schools</td>
<td></td>
<td>n = 11</td>
<td>n = 17</td>
<td>n = 14</td>
<td>N = 14</td>
</tr>
<tr>
<td>bur vs. prof</td>
<td>77</td>
<td>25%</td>
<td>39%</td>
<td>61%</td>
<td>49%</td>
<td>51%</td>
</tr>
<tr>
<td>bur supp prof</td>
<td>26</td>
<td>8%</td>
<td>35%</td>
<td>65%</td>
<td>46%</td>
<td>54%</td>
</tr>
<tr>
<td>prof vs. pol</td>
<td>18</td>
<td>6%</td>
<td>50%</td>
<td>50%</td>
<td>44%</td>
<td>56%</td>
</tr>
<tr>
<td>prof supp pol</td>
<td>19</td>
<td>6%</td>
<td>32%</td>
<td>68%</td>
<td>37%</td>
<td>63%</td>
</tr>
<tr>
<td>market</td>
<td>17</td>
<td>6%</td>
<td>35%</td>
<td>65%</td>
<td>41%</td>
<td>59%</td>
</tr>
<tr>
<td>mor supp prof</td>
<td>16</td>
<td>5%</td>
<td>63%</td>
<td>38%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>bureaucratic</td>
<td>16</td>
<td>5%</td>
<td>25%</td>
<td>75%</td>
<td>88%</td>
<td>13%</td>
</tr>
<tr>
<td>bur vs. mor</td>
<td>15</td>
<td>5%</td>
<td>40%</td>
<td>60%</td>
<td>47%</td>
<td>53%</td>
</tr>
<tr>
<td>professional</td>
<td>15</td>
<td>5%</td>
<td>40%</td>
<td>60%</td>
<td>53%</td>
<td>47%</td>
</tr>
<tr>
<td>political</td>
<td>13</td>
<td>4%</td>
<td>23%</td>
<td>77%</td>
<td>54%</td>
<td>46%</td>
</tr>
<tr>
<td>prof vs. mark</td>
<td>10</td>
<td>3%</td>
<td>60%</td>
<td>40%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>bur vs. pol</td>
<td>9</td>
<td>3%</td>
<td>67%</td>
<td>33%</td>
<td>33%</td>
<td>67%</td>
</tr>
<tr>
<td>bur vs. prof supp pol</td>
<td>8</td>
<td>3%</td>
<td>63%</td>
<td>38%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>bur vs. prof vs. pol</td>
<td>8</td>
<td>3%</td>
<td>25%</td>
<td>75%</td>
<td>63%</td>
<td>38%</td>
</tr>
<tr>
<td>prof supp mark</td>
<td>8</td>
<td>3%</td>
<td>38%</td>
<td>63%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>bur vs. prof supp mor</td>
<td>7</td>
<td>2%</td>
<td>43%</td>
<td>57%</td>
<td>71%</td>
<td>29%</td>
</tr>
<tr>
<td>bur vs. prof vs. mor</td>
<td>6</td>
<td>2%</td>
<td>0%</td>
<td>100%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>bur vs. mark</td>
<td>5</td>
<td>2%</td>
<td>80%</td>
<td>20%</td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td>pol vs. prof suppbur</td>
<td>5</td>
<td>2%</td>
<td>80%</td>
<td>20%</td>
<td>40%</td>
<td>60%</td>
</tr>
<tr>
<td>bur vs. prof vs. mark</td>
<td>4</td>
<td>1%</td>
<td>25%</td>
<td>75%</td>
<td>25%</td>
<td>75%</td>
</tr>
<tr>
<td>moral</td>
<td>4</td>
<td>1%</td>
<td>50%</td>
<td>50%</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>Total/Average</td>
<td>306</td>
<td>100%</td>
<td>43%</td>
<td>57%</td>
<td>52%</td>
<td>48%</td>
</tr>
</tbody>
</table>

Note. This chart represents only the top 91% of incidents. When coded, the remaining 9% of incidents each represented less than 1% of the total, which we take to reflect idiosyncrasies in either coding or principal experiences of accountability.
comparisons (columns 8–13) show how principals constructed their accountability environments when sorted by our measures of the educational task: the diversity index and the peer index.

Gender

Principals in both our samples were about 40% women; relatively high compared to a national average in 2000 of 21% women high school principals (Gates et al., 2003, p. 78). Although old stereotypes are shifting (Glasser & Smith, 2008; Hyde, 2005), much research on leadership still concludes that women, on average, lead in a more democratic or participative manner and engage in more contingent reward behaviors than men. They are rated less effective if they are directive, a putatively masculine behavior, and more effective where the organizational culture encourages communal interpersonal skills (Brunner & Schumaker, 1998; Hoyt, 2007). Consequently, we anticipated finding more female than male principals reporting moral and political accountabilities—which draw on empathy, collaboration, and interpersonal skills—to be critical to their leadership responsibilities. We also anticipated women would report less conflict than men.

We were wrong. Female principals identified political accountability as an influence less often than male principals did, and moral accountability by itself was mentioned no more often by female than it was by male principals. Instead, three unanticipated differences surfaced in which conflict outweighed alignment. Female principals described their moral compunctions as aligned with their professional obligations, but they also saw their professional and political obligations in conflict with bureaucratic ones, and their professional and bureaucratic obligations in conflict with market pressures. All of these types of incidents were mentioned more than we would have predicted given a 40–60 gender split. However, none represented a significant difference in how female and male principals enacted their accountability environments.

School Size

We defined small high schools as those with 550 or fewer students.11 Small high schools are a trend: In 1995–96, New York City had about 60, compared to 99 large ones; by 2004–05 the number was about 160, alongside 121 large schools. Three years later, the number of small high schools was nearly double that of large ones: 205 to 114 (New York City Department of Education, 2008b).

Intimacy is one of the intended effects of small schools. It encourages adult collaboration and nurtures students who need individualized support (Ancess, 2003). But the benefits of intimacy reliably occur only when schools adopt intensely collaborative learning structures (Lee & Smith, 1994), hire teachers who fit these cooperative expectations (Darling-Hammond, Ancess, & Ort, 2002), and when students experience consistency from adults (Lee & Burkham, 2003). There are also negative effects of small school intimacy: “Proximity may breed conflict, and very strong interpersonal relationships may have a negative influence on members’ decisions and behaviors”

---

11There is debate about how to define small schools. Some set the cut-point at 600, others at 500 (see Honig, 2009/this issue).
(Lee, Smerdon, Alfeld-Liro, & Brown, 2000, p. 164), problems that could be exacerbated in New York City where small high schools frequently share a building. Small schools are also many times more likely than large schools to fall into the extreme categories of high- or low-test score performance, or to report changes in score means and gains from year to year. Although some small-school principals are aware of the problem (Hernandez, 2008a), others may jump to “unwarranted conclusions” about their effectiveness, straining the personal relationships needed to sustain collaborative teaching strategies and individuation (Kane & Staiger, 2002, pp. 95, 99, 102).

Again, our findings were only weakly supportive of our initial hunches. We had anticipated that small–high school principals would report high levels of political or market accountability from parents and external organizations because their intimacy encouraged greater contact. We also expected small-school leaders to report higher levels of professional accountability because of collaborative decision making. We thought small-school principals would report that their professional, political, and market accountabilities were in conflict with their bureaucratic obligations more often than large-school principals did. Yet uncontested bureaucratic accountability influenced small-school principals’ decision making seven times more often than large-school principals’, and small-school principals reported accountability conflicts between bureaucratic and political accountabilities about half as often as did large-school principals. Surprisingly, only small-school principals mentioned uncontested moral accountability as an influence on their decisions. Overall, we are left with findings that suggest neither gender nor school size influenced our principals’ accountability environments in ways that the literature led us to expect.

School Diversity

To account for the longstanding research finding that student performance is predicated on student socioeconomic characteristics (Teddlie et al., 2000), we adopted as one measure of the school’s educational task RAND’s school diversity index: “the probability that two members [students] of the [school] population chosen at random will be of different subpopulations” (White, as cited in Gates et al., 2003, p. 75). In RAND’s study of national survey data, diversity values above .7 were associated with a marked increase in principals’ perceptions of poverty-related instructional problems and conflict among parents and community members (p. 123). Because our sample is small, we categorized the RAND scale. Our diversity index scores ranged from high (.7–1.0), to moderate (.4–.6), or low (0–.3).

High diversity schools do not appear to be more conflictual accountability environments for principals in New York City, although we anticipated that we would see results similar to the RAND study. Instead, we saw slightly less conflict in the most heterogeneous schools. Rather than finding higher amounts of conflict between parental demands (whether political or market) and principals’ professional obligations in high diversity schools, as we anticipated, they experienced just the opposite: Professional and political accountabilities supported one another (were aligned) almost 3 times as much as in lower diversity schools. More in keeping with our expectations, high-diversity school principals also reported that accountabilities were in three-way tension many

---

12 A value of zero indicates complete homogeneity; all students are of one racial or ethnic group. A value of one means complete heterogeneity; all groups are represented equally.
more times than did principals in lower diversity schools; principals at high-diversity schools were the only ones to report three-way conflict between market, professional, and bureaucratic accountabilities.

Student's Prior Performance

The Peer Index is New York City's quantification of the educational challenge each high school faces. It is a weighted average of the students' eighth-grade English-Language Arts and math exam rankings that range from 1 to 4.5; the higher the number, the higher scoring the school's entering students. Research using prior achievement as a baseline is intended to evaluate the relative success of interventions by controlling for what the students already know. The difference between where they started and where they end up is measured as a gain (see Roderick, Jacob, & Bryk, 2002). New York City's DOE argues that the peer index is the best predictor of future student performance because it provides a baseline for judging schools on a growth model. Our objective in using the peer index is to acknowledge that principals in low performing schools face a difficult educational challenge: If students are to graduate, schools must focus on accelerating student performance. Consequently we anticipated that principals of low peer index schools would report lower levels of conflict between their internal moral and professional obligations and the political or market accountabilities they have to parents and other community constituents than did principals of high peer index schools.

Our expectations were largely confirmed. Although principals of schools with lower achieving students told us about complex conflict situations more often than principals of higher achieving schools, they also told us that their professional and market accountabilities were aligned; out-distancing high-achieving schools by a factor of 2 to 1. Lower achieving school principals also reported much less conflict between professional and political accountabilities and, in general, reported fewer complex accountability conflicts.

Even so, the reasons for these results may not be straightforward differences in the school's educational challenge. New York City complicated the peer index as a measure of the school's educational task by using it to develop a "peer group" against which the school (and consequently the principal) is rated for purposes of deriving a yearly school grade. Because this difference occurred between 2004 and 2007, it is therefore also captured in between-sample comparisons.

Our findings on measures of the educational task were mixed. The school's diversity seemed not to affect how principals enacted accountability in ways we predicted, and where our hunches about students' prior achievement was sustained, we could attribute it to policy changes. It is to our most striking policy findings that we now turn.

13 The mean peer index score for all 228 high schools that were rated in 2006-07 was 2.73, the median index score is 2.64, the lowest score was 2.04, and the highest was 4.06.

14 Those with lower peer index scores also reported more uncontested bureaucratic and political accountabilities than did their high-peer-index colleagues, likely a reflection of the sampling frame.

15 Each school's peer index is used to identify the 20 high schools whose incoming students score just below, and the 20 schools whose incoming students scored just above it on the eighth-grade tests. Schools facing similar student challenge, it is argued, should be spurred to action by being rated against one another.
POLICY EFFECTS ON PRINCIPALS' SENSE OF THEIR POSITIONAL AUTHORITY

New York City's school accountability regime is an extreme example of an incentive-laden "recognition and reward program" (Ladd, 1996, chap. 2). This form of outcomes-based bureaucratic accountability is intended to reinforce the principal's managerial authority (referred to as "empowering") and is largely drawn from corporate experience. It rests on the coordinated use of regular public reports of school performance against targeted outcomes and personal consequences for principals. One assumption behind increased sanctions and rewards is that principals will figure out how to align school resources to facilitate performance beyond expectations, sometimes referred to as "transformational leadership" (Northouse, 2006). When principals embrace the performance criteria of the district as official protection for decisions they want to make but would otherwise avoid, they become less conflict averse. If this occurs, the principal's positional authority will be experienced as having been enhanced.

We anticipated, therefore, that the adoption of personal sanctions that applied to principals in 2007-08—but not in 2004-05—affect how they experienced their own sense of authority. We expected that 2007-08 principals would attend less to professional and moral accountability in their decision making, and more to bureaucratic, market, and professional accountabilities. We also expected them to identify their accountability environments as more aligned than 3 years earlier, when personnel and curricular decisions were independent of published performance ratings and personal consequences.

Nearly all our hunches about shifts in the principal's authority seem to be born out, and the numerical differences are big. Even so, we do not want to overstate these findings—they are based on relatively small samples and there are many potentially confounding interpretations—so we use the principals' own words to clarify what we found.

MAKING SENSE OF OUR FINDINGS

Figure 1 shows that principals in 2004-05 identified professional and bureaucratic accountabilities as most central to their decision making, and the pattern remained 3 years later. Although some 2004-05 principals reported pressure from the DOE, parents, community groups, and competition (political and market accountabilities), moral sources of accountability were fully one quarter of all incidents. By 2007-08 political and market accountability had replaced some of the professional and moral compunction.

Figure 2 shows another difference in principals' enactment of their accountability environments. Both complex conflict and simple conflict were less evident in 2007-08, replaced by a

16 New York City's reforms were explicitly intended as an extension of Chicago's, where the outcomes-based bureaucratic accountability scheme—also implemented when a mayor took over the schools—was most strongly influenced by state and local corporate business associations (Shipps, 2006).

17 Because the number of small high schools is growing, and our second sample contained some new schools, it may be that our small school measure also interacts with policy changes in accountability. We find this especially plausible since distinctions between small and large high schools showed a pattern similar to that reported in this section.
sharp increase in principals’ perceptions that every type of accountability influenced them directly and without contestation. Table 4 and the principals’ own words help us explore the details.

Enduring Conflicts: Professionalism and Bureaucracy

The number of incidents characterized by bureaucratic and professional conflict tops all others in both periods. All 28 principals told us about at least one such incident, and 21 gave us more than one. Consequently, we divided this huge category into several subcategories to learn more.

Principals in 2004–05 thought of this core conflict in two ways. One was the emerging influence of tested outcomes competing with the pedagogical goals schools had for their students. Principal Pine experienced the conflict emanating from the state: “We tried to provide services for our students who were special ed[ucation] students whether the services were mandated or not and we had a struggle with the [State Board of] Regents. . . . We were basically told that we
had to start providing services only to certain kids because we were not making legal testing requirements." The other was a tension between school-developed instruction and district-wide curricular mandates. One principal spoke for many: "To maintain a steady vision [means] being in every classroom as much as possible, which became more difficult as the system became more focused on the system, rather than the needs of [school] people."

Three years later similar conflicts remained, albeit only for 39% of the 2007–08 principals. "You have to make decisions between how much test prep you're going to do versus how much instruction you do." Principal Bronte went on to say that by 2007–08 this conflict had become a routine dilemma: "[Now] you balance it. You try to do as much of what you consider authentic education, and do things that you think are worthwhile, and you [also] do test prep." Speaking for nearly a dozen of his colleagues, Principal Tan described how the tensions between the district's curricular expectations and school-developed interventions had been institutionalized into two distinct types of accountability. "The [bureaucratic] accountability is not for my purposes. It's whatever data or statistics they want to show. To me, it's moving students, graduating them and moving them to college. That's what it's all about, and making the right decisions for them along
the way." Principal Steinbeck explained that deciding between them was not always an option. "I have a boy. . . He's now in the supposedly tenth grade [15 years old]. No credits this year. He comes to school late every day with a basketball . . . sometimes I have to kick him out of the buildings at 6:30, 7:00, 7:30, 8:00 o'clock at night. . . . He roams the halls. Doesn't go to class . . . He does not want to be here or we do not have whatever it takes to motivate him. [But] he can't go anywhere [else]. Most of the GED programs or other types of alternative programs begin with children who are 17 years old . . . [and] they have to have X amount of credit."

New Forms of Professional and Bureaucratic Conflict

Principal Eliot and five others added budgets. Cuts—whether routine or unexpected—altering even the best plans for improving student performance. "In the past three years [the DOE] hasn't come close to [estimating] an appropriate number [of students] here at Eliot High School and it's affected us drastically twice, where we were short many kids and lost hundreds of thousands of dollars in the middle of the year. Just this past month I lost money because we didn't get enough kids and because the mayor cut everyone's budget. . . . I'm constantly fighting with [and] debating the budget office, [saying] 'I need money for this . . .' And sometimes it doesn't happen. That's when you bang your head against the wall."

The conflict between DOE priorities and principals' professional expectations could become especially frustrating when, as Principal Proust put it, "The Department of Education has made no choices. It's just been, 'Let's do everything.'" As a small-school principal she was certain: "I can't do fifteen initiatives at one time because they're not gonna work, because people are gonna be confused. They feel like this administration, I think, does have lot of good ideas, [but] unfortunately is trying to do, probably, a lot in a little bit of time, before they are gonna be ousted from power. And maybe if they knew that they had a longer timeline, they would be doing it more sensibly." Some felt this tension acutely because their hard-won professional experience was discounted if it predated mayoral control of the schools. Principal Morrison explained that because "the job has changed immensely" she lacks control over "the real things that impact the school:" She asks herself, "Would I do it now, over again, knowing what I know now? I don't think so."

Growing Alignment Between Professional and Bureaucratic Expectations

Reflecting one of the largest shifts we saw—from 20% of incidents in 2004-05 to 72% in 2007-08—principals told us critical incidents in which bureaucratic mandates from the DOE supported their professional judgments. This change reflects specific support structures: the mandate to create a data inquiry team in every school, the SQR process, and unanticipated favors from individual district administrators. Together these accounted for 57% of all such incidents in 2007-08.

Principal Cather exemplified those who singled out the data inquiry team process. "I love the idea that teachers are researchers, because isn't that what your goal is? Let's start with an open-ended [question] and then let's see if we can develop some practices that work." The inquiry team is assembled from among the school's teachers, staff and administrators by the
principal and charged with identifying a small subset of students, about 15, facing serious problems with performance. It is charged with using the scientific method—examining students’ data, hypothesizing interventions to accelerate their learning, conducting the interventions, and monitoring the students’ performance—to improve performance. As Principal Cather described it, “My inquiry team took 15 students, ninth and tenth graders who scored at a level one or two coming from the eighth grade, who seemed to have a problem with fluency. We gave them a series of standardized tests, pretests. And then we moved them along and now what we’re doing is we’re gonna [look at] January scores in English and global studies.” He, too, is learning how to predict student performance: “My first thing was ‘Let’s see the kids who failed global studies in the ninth grade as a predictor.’ [But] there was no correlation ... so we had to look beyond”.

Principal Oates was equally direct about how the SQR process aligned with his professional goals. “We find the quality review self-evaluation, which takes place prior to the quality review, to be one of the most important things that we do. ... My three assistant principals and three or four other teachers actually sit in a room. ... We are brutally honest with ourselves. ... Every time we’re self critical, somebody ... generally me, takes umbrage. ... ‘If you think there’s something we’re doing that doesn’t make us the best at what we’re doing, what can we do to make us better?’” In contrast to the preparation process, the quality reviewer’s rating “is either going to verify what we said is right, or ... is going to give us a higher grade than what we thought we should have gotten, in which case I say ‘Thank you’ and move on.”

Principal Eliot gave us one typical example of unanticipated support from a district administrator that garnered his school an award-winning program. “[The superintendent] spoke to a woman who was running the program for him. ... She had been here and visited us ... And she suggested us. ... She said he turned to her in the elevator and said ‘Yes. Let’s put it in that school. That’s a good school.’ ... And just that conversation ... has dramatically altered the lives of some of these kids.”

Parent’s Altered Influence

In 2004–05 conflicts between the principal’s sense of professional accountability and her political obligations to the school’s external stakeholders, usually parents, were common. Then, principals told stories about parental influence that directly conflicted with their educational goals: For example, the pressure exerted by parents to change the schools’ grading policy: “What your dealing with is parents who went to school and therefore [think they know] what a school is about.”

Notwithstanding the ubiquity of conflicts between school goals and individual parents’ expectations, 3 years later, political accountability was described differently. Several high school principals told us stories about using political negotiation and coalition building skills to engage parents, often by initiating school meetings to discuss problems or establishing routine communication via the internet. Principal Hemmingway’s incident was typical. “Last week, I got an e-mail from ... the PTA president, which I found very disturbing. [He] basically questioned my commitment to the school; and it was based upon misinformation. [Later] when I saw him ... I took out the data for this school for the period of time before I was the principal, and I took out the data for the school during the time for which I was the principal. And I talked about the progress we’ve made and I talked about things that we’ve done to bring ... us that progress. And
he was convinced.” Afterwards, Hemmingway reflected, “I think that to be a principal here, right now, you have to be ... highly resilient. You have to be able to take criticism and respond to it appropriately.”

In 2007–08 more principals reported that political accountability supported their professional decision making if they expanded their definition of constituents to include community-based organizations, businesses and local politicians. “Sometimes when you have a vision and you think, ‘I don’t know how I’m gonna do it,’” Principal Salinger reflected, “you really gotta reach out to people and communicate, because you’d be surprised how willing people are to work with you. Sometimes [principals] just don’t do it. It is hard work; it’s a lot of meetings. It’s a lot of networking, it’s a lot of begging. ... These people [who helped with student internships] are happy that I called them.”

Principal Marquez epitomized those principals who had learned how to reach out to politicians. When still the assistant principal at her school, she had developed an accelerated program for ninth graders that allowed them to graduate early if they attended a longer school day. But the program needed a building. She “had to go to every single community meeting” where she would “talk up the need for it.” She also testified before the Borough president and in the city council. Since then, these politicians recognize her, and several agreed to be guest speakers at events for the school. “Politicians all recognize the need. ... They are all interested in the community.” But Principal Angelou found that supportive political patrons could also engender conflict with the DOE. She explained, “You know what happens when the aunt that everybody’s afraid of comes to town and favors the stepchild? After that aunt leaves, that stepchild gets beat.”

**Competition Becomes a Resource**

In 2004–05, market accountability showed up when in conflict with professional accountability. These incidents were relayed as feelings of “pressure” to recruit quality teachers in competition with districts outside of the city: “Long Island is literally on my doorstep [offering] higher pay, smaller class sizes. They are competing for the very same teachers and university graduates that I am, and I’ve lost.”

When asked if she ever experienced market accountability in 2007–08, Principal Proust’s immediate response was, “We aggressively go after teachers. ... We spend a ton of money and effort on teacher recruitment, everything from having fancy materials that make our school look better than any other school, and banners. ... My hiring coordinator, you know, emails every [education] professor ... at every major university around the country. ... We’re in contact with the Black Students’ Association.” In her view, the school’s competitive advantages—teamwork, a collaborative culture, plus “autonomy in terms of teaching”— did not obviate the need to seek out new teachers each year.

Principals told us more often about harnessing competition to leverage their school’s instructional success by seeking out the “right” students. “I’m a competitive person by nature ... so I think by nature you are always looking at where you stand in comparison to those in your community.” Principal Oates went on to explain how he competed for freshmen. “I study the ones who don’t come. ... [Ours] isn’t always their first choice. There are two other choices they [tend to] make. ... So generally we go through all the applications. ... I get a thousand applicants for seventy-five seats. I’m never in a situation where I go begging for kids, I just need the right
kids.” Another principal had a mental list of three schools against which he ranked himself each time a progress report came out. His list had nothing to do with the peer group established by the Department of Education. Parents, he reasoned, looked at their child’s choices differently than did the DOE.

Principal Eliot encouraged competition within the school. “I give out T-shirts. It started with the students. We have a positive behavior program. . . . And teachers would come to me and say ‘I want one for this kid who did this.’” But he insisted that T-shirts only be a reward for results, “not for each kid, [rather] the kid who had the highest grade in each class.” Those principals who embraced market accountability often said something like, “I find more competition in this, driving this, than I do the accountability from the [DOE] . . . because I know that if I do this, I will never have to sit there and go through accountability checklists.”

Moral Commitments Take a Back Seat

Principals in 2004–05 told us that they felt their moral and professional accountabilities were aligned four times more often than did principals in 2007–08. How principals enacted moral accountability also changed: In 2004–05 it was enacted as a bulwark supporting the principal’s professional commitments, whereas 3 years later we heard much less confidence about acting on one’s value commitments. For example, Principal Birch revealed that her ethical commitment to social justice buttressed her professional commitments to progressive pedagogy, a constant motivation over her 20 years as an educator. “I became a teacher to empower those [students and families] who maybe didn’t have some access . . . we need [access] now more than 20 years ago.” Principal Hemmingway described a different relationship in 2007–08. “Sometimes there is actually a conflict between the two, moral accountability and bureaucratic accountability.” In situations like deciding whether to punish a student infraction by the rules, or bend them to meet circumstances, “I guess which way you go depends upon . . . what level of risk you feel comfortable with.”

Another form of conflict between moral commitments and bureaucratic requirements, common in 2007–08, would not have been possible in 2004–05. Most principals had adjusted their expectations to suit bureaucratic, political, and market evaluations of their professionalism by 2007–08, but they sensed there was something missing. As Principal Morrison described it, “The social growth of students is not . . . evaluated well enough. I don’t know if there is a mechanism for doing it. . . . But to me there has to be a way of measuring a kid that comes to the school in the ninth grade who may have been left back and is reading on third grade level and is living in a shelter, Okay? . . . Some of the kids who did not graduate last year I was very close to, and I saw them as ninth graders as squirrelly, troubled kids, who, granted, did not graduate in four years, but had evolved into a semblance of a functioning adult who could rationalize and reason four years later . . . and these are kids that are riddled with so many problems of society. . . . And to me some of that has to be factored into this scenario somehow.”

Principal Oates mused about moral compunctions being a vanishing perspective. “I, as most people who have worked in the system as long as I have, do become resentful having gone through what I think is the only way to go through the system . . . bring[ing] . . . the human point of view to the principalship. And I worry that the [moral] accountability . . . will be something that some of the people that come after me won’t think is the most important thing.”
REFLECTIONS ON PRINCIPAL POWER

In 2004–05 our principals appeared focused on the school's internal instructional environment and school relationships defined by value commitments. They told us about critical incidents that involved conflicts between external types of accountability—political, bureaucratic, and market—and their professional and moral commitments. They experienced alignment as their professional training reinforced their personal values, and vice versa. Their authority lay in the ability to nurture, protect, and buffer a carefully crafted internal school environment from external demands (Malen & Cochran, 2008).

By 2007–08 our principals were externally focused. All five types of accountability became distinct, uncontested forces shaping their decisions. Instead of dichotomizing internal and external accountabilities, they constructed their professional commitments to support their external ones. Their value commitments were decoupled from their professional ones, as they attended less to ethical precepts or experienced moral accountability in tension with their other responsibilities. Their authority lay in the ability to embrace multiple data sources, using external constituents and competitive forces to create a coherent story about the school.

These changes appear to trump several other sources of variation among high school principals, including variation based on indices of the school's educational task or aspects of its adult relationships. If we are correct, our study extends Leithwood and Jantzi's (2008) observation that district policies explain a large proportion of the variation in how principals see their own effectiveness by changing what principals attend to in the accountability environment and how they relate one obligation to another.

How can this be so if, as research suggests, officials "cannot mandate what matters" (McLaughlin, as cited in Louis, 2007, p. 108); if school leaders make deliberate decisions about how to react to the incentives and sanctions they face (Honig & Hatch, 2004); and leaders are constrained by their own historically contingent narratives about "our school" (Lemons, Lischei, & Siskin, 2003)? The fact that none of our principal informants in 2007–08, although randomly chosen, had more than 6 years of experience as a principal hints at one reason for the swift change our data depict. It can be a rival hypothesis challenging our conclusion, or as we argue below, a contributing factor to a changing politics of the principalship.

Large numbers of new principals have been recently hired in New York City and anecdotal evidence suggests that many seasoned principals have left, perhaps in reaction to the new consequences assigned to bureaucratic accountability. Many of the newcomers lead small schools with external partners, share space with other schools in the same building, and are assigned ambitious outcome targets for student populations small enough to encourage wide year-to-year performance fluctuations. The professional commitments of these new principals are being shaped in a policy environment that actively rewards attention paid to external accountabilities over internal ones. External accountabilities become strategic resources. Embracing district-mandated tools (e.g., inquiry teams, SQR) helps principals describe how their professional priorities align with district goals. Exploiting an advantage in the competition for scarce human resources replaces moral compunction as a motivation to excel. Mastering political negotiation skills helps ensure the perception of progress, even if some targets are not met. If this narrative is reasonably accurate, it suggests that simultaneous policy changes appear to have altered the politics of the principalship.

Most analysis has framed the politics of the principalship as a series of efforts to contain or manage conflict in the face of mandates from policymakers and occasional demands from
middle-class parents and activist teachers. Our findings suggest that this focus on the differential forms of power that principals wield to control the internal school agenda are too limited. They may miss the ways that external interests and constituencies are now shaping principal decisions, including how those external pressures realign principals’ professional expectations.

Those studying the politics of the principalship have routinely argued that it requires special micropolitical lenses to capture. But if the behaviors we have identified persist and prove generalizable, we may make fewer analytical distinctions. Perhaps principals should be studied like other appointed political leaders, complete with constituencies and special interests to satisfy; self-interested, competitive motivations; and professional standards of practice that may align with their superior’s expectations as well as conflict with them. We might ask, What are the relationships between the private interests influencing policy changes and the public authority vested in the principal? How do principals coopt or cooperate with civic and political leaders? How does competition for scarce human resources in a high-stakes environment reshape principals’ strategic decisions? What is lost if principals no longer expect their professional obligations to spring from their moral commitments?

Mayor Bloomberg recently succeeded in obtaining a one-time exemption from the city’s term limit law and has as much as $100 million to spend on reelection for a third term in 2010. The city’s mayoral control law looks set to be reauthorized with relatively minor changes in 2009. Past may well prove prologue for at least 4 more years, and the politics of the principalship may be shifting before our eyes.

REFERENCES


Honig and Hatch (2004) proposed a continuum in which school leaders' reactions range from attempting to coopt the district’s regulators (i.e., bridging strategies) to ignoring incentives and binding their schools to alternative organizations outside the district (i.e., buffering strategies), whereas most principals negotiate with the district to alter specific demands, add peripheral units to limit the effects of demands on day-to-day practices, or symbolically embrace them without changing how the school functions.


