"Staffing to the Test": Are Today's School Personnel Practices Evidence Based?
Author(s): Lora Cohen-Vogel
Published by: American Educational Research Association
Stable URL: http://www.jstor.org/stable/41413068
Accessed: 12/01/2015 16:05

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.

American Educational Research Association is collaborating with JSTOR to digitize, preserve and extend access to Educational Evaluation and Policy Analysis.
“Staffing to the Test”: Are Today’s School Personnel Practices Evidence Based?

Lora Cohen-Vogel
University of North Carolina at Chapel Hill

Faced with mounting policy pressures from federal and state accountability programs, school leaders are reallocating curricula, time, even diet in an attempt to boost student achievement. To explore whether they are using test score data to reallocate their teacher resources as well, I designed a cross-case, cross-sectional study and explored principals’ reported staffing practices in one higher performing and one lower performing elementary school in each of five Florida school districts. Findings show that school leaders are “staffing to the test” by hiring, moving, and developing teachers in an effort to increase their schools’ overall performance. The paper discusses the implications of evidence-based staffing for policy, practice and future research.

Keywords: accountability, teacher quality, data use, personnel practices

Faced with mounting policy pressures from No Child Left Behind (NCLB, 2002) and state accountability programs, school administrators have begun reallocating their resources in an attempt to improve student achievement. District and school officials increasingly instruct faculty to “teach to the test” by aligning curricula and materials to the content and format of state tests (Diamond & Spillane, 2004; Firestone, Monfils, & Schorr, 2004; Guskey, 2003; Lyons & Algozzine, 2006; Stecher & Barron, 1999). School leaders also schedule to the test, reducing time allocated for lunch, recess, and some untested subjects (Center on Education Policy, 2005; Matthews, 2007; McMurrer, 2007, 2008); reclassify to the test by identifying low-performing students as disabled or Limited English Proficient to exclude them from the testing pool (Allington & McGill-Franzen, 1992; Figlio & Getzer, 2002; Haney, 2000; Heilig & Darling-Hammond, 2008); retain to the test, increasing the likelihood that students not promoted to grade 10 would drop out or “disappear” (Heilig & Darling-Hammond, 2008); discipline to the test by reducing suspension penalties for high achievers and raising them for low achievers during testing periods (Figlio, 2003); and feed to the test by increasing calories on test days (Figlio & Winicki, 2002). As school leaders are reallocating curricula, time, and diet in an effort to raise achievement, are they also reallocating what is arguably their most precious resource: teachers?

Education researchers have come to demonstrate that which principals and parents have long suspected—that teachers represent the single largest school-related influence on student achievement (Sanders & Horn, 1998). Although studies (e.g., Greenwald, Hedges, & Laine, 1996; Hanushek, 1989) show mixed results in terms of the effects of individual teacher characteristics (e.g., education, experience) on achievement, there is wide agreement among researchers who have adopted “pure outcome-based measures” of teacher effectiveness that teaching matters and matters significantly for student success (Goe, 2007). Findings from these teacher value-added studies have consistently revealed substantial variation in teachers’ abilities to boost student achievement (e.g., Jordan, Mendro,
In interpreting the size of the effect, Sanders and Rivers (1996) showed that fifth-grade students in Tennessee assigned to very effective teachers 3 years in a row scored, on average, 50 percentile points higher on the state’s math assessment than students who had 3 years of ineffective teachers (see also Hanushek & Rivkin, 2004; Rivkin, Hanushek, & Kain, 2005).

The teacher effect on student gains may be particularly strong for low-income and low-performing children. In 1992, Hanushek showed that good teachers can elicit an additional year’s worth of achievement from low-income minority students whose average achievement in primary grades is below the national average. That is to say, low-income, lower performing minority students assigned to a teacher at the 95th percentile of the quality distribution gained 1.5 grade level equivalents in a single year, whereas similar students assigned to a teacher at the 5th percentile gained only 0.5 years of grade level equivalency. Like Hanushek (1992), Babu and Mendro (2003) found that “the teacher effect” is especially strong for low achievers. According to the authors, when low-achieving fourth-grade students were assigned to effective teachers 3 years in a row, they were twice as likely to pass a seventh-grade end-of-course mathematics test than low achievers assigned to ineffective teachers. By implication, these studies suggest that school leaders may be able to both improve student outcomes and close stubborn achievement gaps with changes to the ways they staff their schools.

This study begins to lay the groundwork for research around this idea—one I will call “evidence-based staffing.” Here, I investigate whether personnel decisions are being made in an effort to improve student achievement and how, if at all, test score data are being used in teacher staffing. Specifically, I ask, How do principals make teacher staffing decisions? And, what role do student test scores play in hiring, assigning, developing, and dismissing teachers? The term teacher staffing here refers to decisions about who will teach, where they will teach, and what they will teach. To date, relatively few studies have explored administrators’ staffing practices and only one has examined the role that teacher performance—and their students’ test score data, in particular—plays in their decisions. My intention here is not to determine whether evidence-based staffing is “right” or effective; that will be left for future research. Rather, using data from interviews and school and district policy documents, I offer an analysis of the ways school leaders in Florida report using achievement data to hire, assign, develop, and dismiss elementary school teachers.

I argue that comparing administrators’ reports of their staffing decisions has implications not only for schools’ capacity to reform but also for equality of opportunity. Finding that administrators in some elementary schools in Florida report staffing to the test, I suggest that such practices may differ between higher and lower performing schools and rely on the assumption that a “good” teacher is good for all students equally. I argue that these new test-based staffing practices may represent another example of the ways in which schools, under unprecedented demands for achievement and with relatively little empirical evidence about “what works,” have become incubators of new practices, and I caution that unless robust empirical attention is paid to test-based staffing practices, we risk missing an opportunity to understand the effects of innovations brought about by standards and accountability frameworks that currently dominate state and federal efforts for school improvement.

**Teacher Staffing Decisions: What We Know and What We Need to Know**

References to staffing in the literature appear primarily in studies related to teacher supply. These studies focus on factors that lie largely outside the control of individual school administrators—factors like market conditions and teacher preferences about where to teach. With a consistency that is rare in education research, the studies show that these factors produce variation in both the supply and quality of teacher applicants, with schools located in urban and rural settings and enrolling lower performing students, in general, being harder to staff (e.g., Levin, Mulhern, & Schunk, 2005; Liu, Rosenstein, Swan, & Khalil, 2008). Less is known, however, about what economists call the demand side of the equation—the ways administrators weigh decisions about whom among the pool of applicants to hire and where to place them once hired. Recent work has begun to plug those holes, although research on the role that performance...
data play in principals’ staffing decisions remains limited. In the rest of this section, I will review the research on teacher hiring, assignment, development, and dismissal in turn.

**Hiring Practices**

Recently, Rutledge, Harris, Thompson, and Ingle (2008) organized studies on teacher hiring into three groups: those that seek to understand the hiring process, those that consider constraints on teacher hiring, and those that identify administrators’ preferences for applicant characteristics.

**Teacher screening and hiring process.** Research has identified a set of norms that characterize the hiring process, including a summer hiring schedule, a screening stage, and applicant interviews (Rutledge et al., 2008). Within these standards of practice, school districts seem to use different approaches to screen and hire applicants. School districts with “highly decentralized” approaches allow individual schools to accept applications and hire into open positions; in districts with “highly centralized” approaches, district officials oversee screening, selection, and assignment processes. “Moderately centralized/moderately decentralized” approaches are taken in still other districts where district officials screen applicants, and school administrators, with or without teacher input, select from among them (Liu & Johnson, 2006).

Regardless of the approach that districts take, school principals appear to exercise considerable influence over teacher hiring and, in most districts, remain the “primary gatekeepers” for decisions about who teaches in their schools (Rutledge et al., 2008, p. 243). Indeed, principals are involved in teacher hiring decisions more than 75% of the time (Liu, 2002; Strauss, Bowes, Marks, & Plesko, 2000).

**Constraints on teacher hiring.** Principal influence on hiring is not without its limits, however. A handful of recent studies (Ellsasser, 2008; Guin, 2004; Hess & West, 2006; Hill, 2006) blame staffing rules in teachers union contracts for forcing public school systems to hire “teachers in urban classrooms each year with little regard for the appropriateness of the match, the quality of the teacher, or the overall impact on schools” (Levin et al., 2005, p. 7). These seniority rules, according to the authors, mean that administrators in the urban districts studied have little or no choice over the selection of 4 out of every 10 teachers hired; vacancies in these districts are filled instead by existing teachers who request transfers to the school or teachers previously laid off.

Even when school leaders have discretion over teacher selection, they have been criticized for preferring “to hire locally or to favor those most familiar to them” (Baker & Cooper, 2005, p. 450). From a survey of 510 school districts in Pennsylvania, Strauss and colleagues (2000) found that (a) 40% of teachers obtained their high school diplomas or attended high school in districts where they worked; (b) about one third of districts filled full-time openings with substitutes or part-time teachers they already knew; and (c) districts tended to hire teachers trained at local colleges. Whether these findings reflect a bias for locally grown teachers or simply a preference among teachers to work in the cities and towns in which they were reared is difficult to say (Boyd, Lankford, Loeb, & Hanushek, 2005); what they do suggest, however, is that school leaders may not be using hiring practices formulated for finding the very best possible candidates.

**Preferred applicant characteristics.** In hiring from among teacher candidates, principals appear to look for attributes in applicants that complement those of the existing faculty (Abernathy, Forsyth, & Mitchell, 2001; Cain-Caston, 1999; Harris, Rutledge, Ingle, & Thompson, 2010; Theel & Tallierico, 2004). Specifically, “they ‘mix and match’ teacher characteristics to fill important voids in racial and gender diversity, as well as skill, and to ensure that teacher candidates fit the cultural norms and values of the school as an organization” (Rutledge et al., 2008, p. 255). To mix and match, principals use a variety of tools including the face-to-face interview, reviews of certification status, resumes, and transcripts, and recommendations from other principals and teachers (Rutledge et al., 2008).

The reliance on such tools, however, may result in a hiring process that is “information poor” in terms of reliable indicators of teacher effectiveness (Liu & Johnson, 2006, p. 324). Administrators’ preferences for teacher candidates often focus on organizational skills (Harris et al., 2010) and
“rarely are more objective measures of teaching quality, such as student outcome data, included in these analyses” (Baker & Cooper, 2005, p. 451).

A considerable number of studies have looked at SAT scores and/or the selectivity of the undergraduate institution from which teacher candidates graduate as indicators of teacher quality (e.g., Angrist & Guryan, 2004; Ballou & Podgursky, 1995). Drawing on a series of large, nationally representative surveys of new college graduates, Ballou (1996) found that college selectivity has virtually no influence on a candidate’s likelihood of receiving an offer to teach full-time in a public school, although applicants with higher grade point averages were found to be given modest preference. A thorough review of the literature found no studies of whether and how principals value (and use) more direct indicators of teacher quality—namely, student outcome data—when hiring from among teacher candidates. This study is designed to begin to fill that gap.

Assignment Practices

Few studies have considered the ways in which school leaders assign teachers to individual grades and subjects. Those that have do so indirectly are grounded in the tracking literature, and pre-date contemporary accountability policies in education. In a study of student scheduling, for example, DeLany (1991) found that frequent changes in student enrollment, teacher turnover, and revised graduation requirements lead to “shoehorning” wherein high school teachers are reassigned to courses for which they may hold no certification or have not been adequately prepared. As a result, “out-of-field” teaching is widespread, especially in secondary schools where more than one fourth of students enrolled in mathematics classes and nearly one half in life and physical science classes are taught by teachers with no major or minor in the field (Ingersoll, 2003; Ingersoll & Gruber, 1996; Wirt et al., 2004), leading Congress to pass legislation requiring teachers to possess certification in the core subjects they teach.

There is limited evidence that assignment processes may also reflect political priorities. Administrators may match teachers and students to classrooms under both demand- and supply-side pressures (Cohen-Vogel & Osborne-Lampkin, 2007). On the demand side, parents, especially those with time, information, and status, may use voice and the threat of exit to press principals to make assignment decisions that favor their children (Clotfelter, Ladd, & Vigdor, 2006). Other typically less-advantaged parents appear to be more likely to leave decisions about their children’s education to the “professionals” (Smrekar & Cohen-Vogel, 2001; Lareau, 1987). On the supply side, teachers articulate their preferences for assignments to particular subjects, students, and grades and use various strategies to avoid teaching lower tracked courses. In a case study of one suburban high school, Finley (1984) found that teachers create new courses “which, by virtue of their content and the way they have been advertised, have attracted the students these teachers want to teach” (p. 238). They also intentionally develop reputations as being too hard on lower track students and may encourage low performers to transfer to another course during the first few weeks of school. In general, these dual forces—unequal parent pressures and teacher preferences for higher track assignments—appear to result in assignments that pair high-quality teachers with higher performing students (see, for example, Gamoran, Nystrand, Berends, & LePore, 1995; Gamoran, Porter, Smithson, & White, 1997).

Like critics of teacher hiring practices, some have argued that provisions in collective bargaining agreements may also lead school leaders to place teachers in classrooms without consideration for the appropriateness of the match. Bargaining agreements between unions and management in some school districts permit senior teachers to self-select into positions (e.g., subjects) as long as they hold the appropriate certifications (Ballou, 2000; Johnson & Donaldson, 2006). In many other school systems, depending in part on the state in which they sit, however, principals make assignment changes freely without having to weigh teacher seniority (Cohen-Vogel & Osborne-Lampkin, 2007). But, little has been written about the criteria they do use in their decisions and what role, if any, pressure to raise student test scores and performance data play in that process.

Professional Development

In the United States, “planning for professional development activities can occur at any level of
the formal school district hierarchy” (Desimone, Porter, Birman, Garet, & Yoon, 2002, p. 1273). Sometimes, district administrators plan professional development (PD) activities for teachers; at other times, PD planning occurs at the school level, by full faculties, by groups of teachers, or by school leaders with or without teacher involvement (Sykes, 1996). As a result, schools often do not have a coherent, coordinated approach to PD and instruction, and the quality of PD activities varies widely (Desimone, Porter, Garet, Yoon, & Birman, 2005; Garet et al., 1999; Garet, Porter, Desimone, Birman, & Yoon, 2001).

As far back as the early 1990s, scholars recognized the potential for standards and assessment systems to improve PD coordination and quality (Cohen, 1995; Cohen & Hill, 2000, 2001). Others argued for the critical role that PD could play in mediating the success of standards-based reforms through instructional improvement (Cohen & Spillane, 1992; Smith & O’Day, 1991; Spillane & Jennings, 1997). Because they often set the context for PD activities (Elmore & Burney, 1996; Knapp, Zuck, Adelman, & St. John, 1991; Spillane & Jennings, 1997; Spillane & Thompson, 1997), districts were seen as key players for designing PD opportunities in ways that supported state instructional policy (Elmore & Burney, 1999; Goertz, Massell, & Chun, 1998; Spillane, 1996). Scholars argued that in providing high-quality PD, districts must develop and use performance indicators to assess needs, set forth goals, and establish benchmarks for progress monitoring (Fuhrman, Clune, & Elmore, 1988; Guskey, 1997; Loucks-Horsley, Hewson, Love, & Stiles, 1998).

Yet, in 2002, Desimone et al. reported that few districts were using performance indicators in the design and development of PD. Due to a lack of empirical evidence, there are open questions around whether substantive changes have emerged in the years since NCLB. Given the prescriptive nature of the legislation—particularly with regard to schools whose students have not made Adequate Yearly Progress (AYP), wherein, for example, at least 10% of Title I funds must be spent on PD linked to instruction of student subgroups not making AYP—we might predict that both districts and schools have begun to use student outcome data to plan for schoolwide and individual teacher PD.

**Dismissal**

What does the empirical literature tell us about the criteria and data used by school leaders in their decisions to dismiss practicing teachers? The answer is “very little”—probably because teacher dismissals are rare. Indeed, according to Tucker (1997) and others (Bridges & Gumport, 1984; Fuhr, 1993; Lavelly, Berger, & Pollman, 1992), the teacher dismissal rate is substantially less than 1%—a rate far below the 5% to 15% estimates of teacher incompetence (Bridges, 1992; Groves, 1986; U.S. Department of Education, 1993). Recent evidence suggests that of teachers who are dismissed, very few are terminated for poor performance even as school leaders may counsel ineffective teachers out of the profession (Weisberg, Sexton, Mulhern, & Keeling, 2009).

The reasons cited for the low incidence of teacher dismissals are manifold. First, some districts and schools suffer from critical shortages of teachers; terminating teachers in those contexts is difficult. Others blame an onerous dismissal process and grievance and arbitration provisions in collective bargaining agreements that burden principals with years of time-consuming evaluation and paperwork (Johnson & Donaldson, 2006). Researchers have also cited a lack of “resolve” on the part of school principals as a reason for the low rate of dismissals (Phillips & Young, 1997; Pratt, 1996; Randklev & Lemon, 1990). Some principals report feeling unsure in their knowledge about the dismissal process; others say they lack the skills to navigate dismissal proceedings effectively (Hess & Kelly, 2007; VanSciver, 1990).

There are early indications that dismissals for performance reasons may be on the rise; changes in local evaluation practices and federal accountability pressures may be responsible. For example, in one California district, after the implementation of peer assistance and review, in which lead teachers conducted personnel evaluations of 88 beginning teachers and 3 veteran teachers who had received unsatisfactory evaluations the year before, 12.5% of the new teachers were nonrenewed for employment during the program’s first year and all 3 veteran teachers were encouraged to retire, a major shift from prior dismissal rates in the district (Goldstein, 2007, 2009). Indeed, “the union president could not
recall (and the district had no record of) any teachers being dismissed for issues of teaching quality in the years prior to the [program] (Goldstein, 2007, p. 496). In terms of federal accountability pressures, districts across the nation who want a share of the $3.5 billion in School Improvement Grants from Washington must weigh options for turning around their failing schools that include extending instructional hours, converting them to charters, or replacing the principal and at least half the staff. In a 5–2 vote on February 23, 2010, and later endorsed by President Obama, the school board in Central Falls, Rhode Island, approved a plan by its superintendent to discharge the principal and all 93 teachers and staff members of Central Falls High School (Greenhouse & Dillon, 2010).

Although teacher dismissals for performance reasons may be increasing, relatively little is known about how student achievement data are weighed in decisions to dismiss. There is one notable exception. Jacob (2010) examined how principals in Chicago made dismissal decisions after a new collective bargaining agreement was signed in 2004 that gave principals the flexibility to dismiss probationary teachers for any reason and without documentation. He matched information on teachers eligible for dismissal with records indicating which teachers were dismissed and found that principals do consider value-added achievement measures, along with teacher absences and several demographic characteristics, in determining which teachers to dismiss. The remaining evidence in the literature focuses not on teacher dismissals but on their corollary: teacher evaluations. In a longitudinal analysis of teacher evaluation statutes and department of education regulations, Hazi and Rucinski (2009) traced changes to states’ evaluation policies since NCLB. They found that the majority of the nation’s 50 states have increased oversight of local teacher evaluation practices, decreased the frequency of veteran teacher evaluation, and—most notably for this study—strengthened the reliance on student achievement data in teacher evaluations.

At the school and district levels, only a handful of empirical studies in the educational policy and leadership literatures have examined “data-driven decisionmaking” (or D3M) (see, for example, Halverson, Grigg, Prichett, & Thomas, 2005; Wohlstetter, Datnow, & Park, 2008), and only one directly addressed the use of student achievement data in teacher evaluation and dismissal. School administrators have written in practitioner journals, however, about their personal experiences with D3M. Cohen (2003), for example, discussed the use of test score data in teacher evaluations and PD planning in his New York district: “If we see that a particular teacher has average students for three consecutive years who perform below their classmates, we can conclude that the teacher’s effectiveness is below average, allowing supervisors to offer assistance where it is most needed.” He goes on to advise that “before using data analysis to make judgments about a teacher’s effectiveness, it is important to review . . . class enrollment, student [composition] . . . and previous performance by these students” (p. 54). This study will begin to look systematically at these practices, exploring the ways administrators at the school level use test score data in their evaluation and dismissal decisions.

Methodology

To answer what teacher staffing practices look like today and what role, if any, student achievement data play in hiring, assigning, developing, and dismissing public school teachers, I used a cross-case, cross-sectional study within a common state policy context. I explored teacher staffing practices in two elementary schools in each of five Florida school districts. By holding constant state policies that may affect teacher staffing patterns (e.g., class size reduction; tenure laws; due process protections), single-state studies represent important tools for exploring at once multiple, competing explanations for variation in assignment practices between districts, even as they limit generalizability to other states. Because the study included multiple districts, I was able to consider how embedded district contexts and procedures shape administrators’ decisions about teacher hiring, assignment, and dismissal. Involving more than one school in each of the districts allowed me to begin to compare staffing practices between schools with different performance histories.

Because I initially expected staffing decisions to be conditioned by district–union relationships and the collective bargaining agreements between them, a stratified random sampling framework
was used based on contractual flexibility and enrollment to select the districts. The theoretical literature, along with a handful of empirical studies, suggests that union power, and the provisions of collective bargaining agreements, in particular, may condition decisions about which teacher applicants to hire, where to place them, and, to a somewhat lesser degree, to what professional development activities they should be directed (see, for example, Ballou, 2000; Hess & West, 2006; Lieberman, 2000). [As I report elsewhere (Cohen-Vogel, 2008), little evidence suggests that this is the case in Florida.] With a graduate student, I content analyzed collective bargaining agreements in 66 of Florida’s 67 school districts (see Cohen-Vogel & Osborne-Lampkin, 2007). Contract data were analyzed for frequencies (e.g., references to seniority), classified by content area (e.g., new hires, teacher transfers), and coded (e.g., level and type of administrative discretion). Information from the coded data was used to categorize districts by their agreements with teachers unions on a 3-point scale from “more restrictive” to “restrictive” to “less restrictive.” District enrollment data (small, mid-sized, large) were added with the expectation that enrollment may affect assignment practices inasmuch as it enhances or constrains administrators’ staffing options. That is to say, administrators in districts with fewer students and, therefore, fewer schools might be limited in their ability to transfer teachers to other schools. One small, one mid-sized, and one large district from the more restrictive and less restrictive categories were selected at random to make up the district sample; one district was removed from this analysis when I was unable to recruit schools with different demographics and performance statuses to participate.

Performance and demographic data were collected for all non-charter elementary schools in the sampled districts. Specifically, I obtained school grades as well as percentages of minority and English language learner (ELL) students and students eligible for free and reduced priced lunch. Since 1999, Florida has used a comprehensive system of school grades to rate each school on an A through F scale. The primary

---

TABLE 1
Demographic and Performance Profile of Participating Schools

<table>
<thead>
<tr>
<th>School</th>
<th>School grade</th>
<th>5 year school grade trajectory</th>
<th>Enrollment</th>
<th>% FRPL</th>
<th>% minority</th>
<th>% ELL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrus Elementary</td>
<td>A</td>
<td>steady</td>
<td>&lt;400</td>
<td>35</td>
<td>20-25</td>
<td>3</td>
</tr>
<tr>
<td>Lafayette Elementary</td>
<td>B*</td>
<td>up</td>
<td>700-800</td>
<td>50</td>
<td>5-10</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Pine Grove Elementary</td>
<td>A</td>
<td>steady</td>
<td>700-800</td>
<td>15</td>
<td>10-15</td>
<td>2</td>
</tr>
<tr>
<td>Orange Elementary</td>
<td>B*</td>
<td>mixed</td>
<td>700-800</td>
<td>70</td>
<td>30-35</td>
<td>1</td>
</tr>
<tr>
<td>De Maude Elementary</td>
<td>A</td>
<td>up</td>
<td>700-800</td>
<td>80</td>
<td>90-95</td>
<td>60</td>
</tr>
<tr>
<td>Millstone Elementary</td>
<td>C*</td>
<td>down</td>
<td>600-700</td>
<td>85</td>
<td>85-90</td>
<td>55</td>
</tr>
<tr>
<td>Armstrong Elementary</td>
<td>B*</td>
<td>up</td>
<td>500-600</td>
<td>80</td>
<td>95-100</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Cocoa Gardens Elementary</td>
<td>F</td>
<td>down</td>
<td>700-800</td>
<td>65</td>
<td>95-100</td>
<td>20</td>
</tr>
<tr>
<td>Norwood Elementary</td>
<td>A</td>
<td>up</td>
<td>600-700</td>
<td>80</td>
<td>95-100</td>
<td>11</td>
</tr>
<tr>
<td>Sweet Bay Elementary</td>
<td>C*</td>
<td>steady</td>
<td>800-900</td>
<td>75</td>
<td>90-95</td>
<td>25</td>
</tr>
<tr>
<td>All schools (actual average)</td>
<td>—</td>
<td>—</td>
<td>789</td>
<td>63.4</td>
<td>65.2</td>
<td>15.9</td>
</tr>
</tbody>
</table>

Note. In order to preserve confidentiality, schools have been given pseudonyms. Data were rounded or presented in ranges to protect the schools’ identities. They were derived from the Florida Schools Indicators and School Grades Reports. Dark rows indicate schools located in districts with more restrictive collective bargaining agreements (CBAs); light rows indicate schools located in districts with less restrictive CBAs. FRPL = free and reduced priced lunch eligible; ELL = English language learner.

ALowest performing elementary school in the district.

BLowest performing elementary school in the district that agreed to participate. During the sample year, there were two lower performing D schools in the district.

HLighest performing elementary school in the district.

ILowest performing elementary school in the district that agreed to participate. During the sample year, there were two D schools and one F school in the district.
criterion in calculating school grades is student scores on the Florida Comprehensive Assessment Test (FCAT). Wherever possible, I randomly selected one “A” school and one of the district’s lowest performing schools to participate. When only one school in a district had earned an A, it was selected; if no elementary school in the district had earned an A, the highest performing school in the district was selected. When only one school fell in the lowest grade category earned by elementary schools in the district, it was selected. The study was limited to elementary schools in order to control for certification effects on staffing decisions. In general, teachers in elementary schools hold general elementary certificates and can teach grades K through 5. See Table 1 for demographic and performance information on the participating schools.

District informants included superintendents, deputy superintendents, and human resource/employee relations officers. Within the 10 schools, interviews were conducted in person with the principal and assistant principals, as appropriate, school union representative, two teachers selected at random, and parents who served in leadership positions in the school. In all, 61 participants were interviewed, including 8 superintendents/deputy superintendents, 5 human resource directors, 15 principals/assistant principals, 27 teachers, 8 of whom were site union representatives, and 6 parents.

In semi-structured interviews ranging from 45 minutes to 2 hours, discussions focused primarily on the hiring, reassignment (within schools), and dismissal processes. Key questions for principal participants included, for example, “Tell me about the process through which teachers are assigned to teach certain grades or subjects in your school,” and “How would you characterize your staffing philosophy?” Interviews with teachers largely focused on their own hiring experiences as well as the assignments they have held. Having set out to understand how unions and the contracts they make with school districts shape staffing practices, I directed a portion of the interview guides deductively at the school–union relationship and the way union contract provisions may condition staffing practices: “In general, does the grievance and arbitration process affect your [the principal’s] decisions about whom to hire? If so, how?” and “Do you feel that you have the flexibility you need to make ‘good’ staffing decisions?” In this sense, the original interview guides combined structured questions with more open-ended items designed to allow participants to introduce new or “emergent” ideas (Whyte, 1982).

Following Strauss (1987) and Corbin and Strauss (2008), I adopted a sequential, some say “iterative,” approach to data collection and analysis, an approach that allows researchers to “identify relevant concepts, follow through on subsequent questions, and listen and observe in more sensitive ways” (p. 57). This theoretical sampling approach enables researchers to discover concepts that are relevant to a particular population or phenomenon and explore them in depth; the purpose is to build understanding, not count occurrences. With this approach, the researcher is sampling concepts and not persons, and concepts derived from the data drive the next round of data collection activities. In line with the approach, I began analyzing data after completing the first series of interviews. In this way, concepts in the data informed data gathering activities at subsequent points in the study, allowing me to revise my interview guides to follow up on important theoretical leads. Those leads, as I describe below, coalesced around motives behind principals’ staffing decisions and highlighted the role that student achievement data play in hiring, assigning, developing, and dismissing teachers.

Pattern coding of interview transcripts and documents was used to reveal central constructs in the data (Fetterman, 1989; Miles & Huberman, 1994; Yin, 1989). I began by analyzing the data categorically, first assigning basic, descriptive codes for the type of staffing decision under discussion: hiring, assignment, professional development, dismissal, and so on. These codes became the units by which the findings in this article are organized. I then identified key motives behind school staffing decisions. Motives included the grounds for actions taken with regard to the hiring, assignment, development, and dismissal of instructional staff and included “improved test performance,” “teacher preferences,” and “path of least resistance,” for example. Iteratively labeling properties and dimensions in the data, I added other codes to capture the information that school leaders use in their staffing decisions: test scores, recommendations, and observations.
Particular attention was paid to the use of performance data in teacher staffing decisions. I also looked for response patterns across informant groups within single schools and districts and between like informants (e.g., principals) among different sites. Codes were added to the data files to capture district size, school enrollment, and school performance on state tests to analyze relationships between these factors and informants’ responses.

I supplemented these interview data with content analysis of district and school documents that included teacher candidate interview forms, personnel policy manuals, annual teacher unit allotments, and assignment preference forms. Although at times in this article I employ data from the larger sample of interviews and documents, the majority of the analysis reported here draws on data from interviews with school-level administrators and, to a lesser extent, teachers. To help establish dependability, I employed systematic, iterative coding approaches (Strauss & Corbin, 1990), “member checked” with informants, and looked for discrepant cases that did not map readily onto the codes and the connections I made between them (Lincoln & Guba, 1985; Miles & Huberman, 1994).

**Findings**

Findings suggest that school leaders in Florida are staffing to the test. Principals describe hiring, developing, and removing teachers in an effort to increase their schools’ overall FCAT performance, and they report paying particular attention to student test scores in their decisions to reassign teachers within their schools. This section describes how principals say they undertake these decisions and the role that achievement data play in them.

**Hiring and Initial Placement**

All principals in the study report using student achievement data in their hiring decisions. But, they also report doing so in a variety of ways. As they reflect on their hiring practices, principals say they use performance data to assess their staffing needs, to choose among applicants for teaching positions, and to place new teachers once they are hired.

In assessing their staffing needs, three principals talk of using test score data to identify grades and subjects where student test scores have foundered. They report making staffing changes and hiring new teachers in an attempt to shore up those areas. In thinking about her hiring decisions, the principal of a lower performing, Title I school whose performance had dropped two letter grades in the past 5 years remarked, “Sometimes it may be experience, and sometimes it may be that I felt [the candidate] had a real true knowledge in an area where we had a need based on the data from my school.” The principal of an “A” school spoke of using test score data to identify his school’s staffing needs. With the school at risk of losing its A status for failing to bring up the scores of the bottom quartile of students, he reported looking for candidates with the drive to work with “below-level” learners:

That’s something I try to get at when I’m hiring: Are [candidates] willing to work with those kids who are not easily taught, and that’s the bottom line with me. As you know, the state has just implemented this year—they’re looking at your lowest 25 percent.

In terms of selecting among candidates to hire, principals use achievement data to identify effective teachers. Teacher-level student outcome data are not available, of course, when principals are reviewing the application files of new teacher candidates. But, findings show that principals in more than half of the sampled schools request the performance histories and/or test score data of applicants with prior teaching experience.

Principals rely primarily on phone calls with applicants’ past principals for this information. As illustrated in Table 2, principals, without exception, report that they do not hire transfer candidates without first speaking with administrators at applicants’ former schools. When asked what information he uses in his hiring decisions, an assistant principal in a Title I, “B” school whose grade has increased from an F in the past 5 years replied,

If they have been teaching prior, then you ask the principal, or whomever the principal designates that you can talk to in reference to the FCAT scores. If it’s a second grade teacher, then it’s the SAT-10 scores... I usually ask for percentages: “How many students scored a level three, four or five?” And then, “How many were not successful?”

Principals say that they are careful not to ask about test scores “in a vacuum;” however. In

---

491
<table>
<thead>
<tr>
<th>School</th>
<th>Calls for recommendations</th>
<th>Makes assignments freely</th>
<th>Estimated # of reassignments per year</th>
<th>Collects teacher assignment preferences</th>
<th>Uses test scores in assignment decisions</th>
<th>Uses test scores in professional development decisions</th>
<th>Has transferred poor performers out</th>
<th># of teachers on improvement plans</th>
<th>Avg. # of nonrenewals per year</th>
<th># of tenured teachers principal has terminated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Citrus</td>
<td>Y</td>
<td>Y</td>
<td>&lt;1</td>
<td>on spring form</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>1</td>
<td>&lt;1</td>
<td>0</td>
</tr>
<tr>
<td>Lafayette</td>
<td>Y</td>
<td>Y</td>
<td>1–2</td>
<td>on spring form</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>3</td>
<td>1–2</td>
<td>0</td>
</tr>
<tr>
<td>Pine Grove</td>
<td>Y</td>
<td>Y</td>
<td>2–3</td>
<td>informal requests</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>0</td>
<td>1–2</td>
<td>0</td>
</tr>
<tr>
<td>Orange</td>
<td>Y</td>
<td>Y</td>
<td>3</td>
<td>informal requests</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>De Maude</td>
<td>Y</td>
<td>Y</td>
<td>3–4</td>
<td>on spring form</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>0</td>
<td>n/a</td>
<td>0</td>
</tr>
<tr>
<td>Millstone</td>
<td>Y</td>
<td>Y</td>
<td>n/a</td>
<td>on spring form</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Armstrong</td>
<td>Y</td>
<td>Y</td>
<td>1–2</td>
<td>informal requests</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>n/a</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Cocoa Gardens</td>
<td>Y</td>
<td>Y</td>
<td>8</td>
<td>informal requests</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>2</td>
<td>4–5</td>
<td>0</td>
</tr>
<tr>
<td>Norwood</td>
<td>Y</td>
<td>Y</td>
<td>regularly</td>
<td>on spring form</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>2</td>
<td>n/a</td>
<td>0</td>
</tr>
<tr>
<td>Sweet Bay</td>
<td>Y</td>
<td>Y</td>
<td>6</td>
<td>on spring form</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>1</td>
<td>&lt;1</td>
<td>0</td>
</tr>
</tbody>
</table>

*Note.* Dark rows indicate higher performing schools; light rows indicate lower performing.
addition to an applicant’s test score data, reported one principal, “I also ask the [referring] principal what was the history or background with that particular school because I really want to know what type of children she was teaching.” Another principal said that, apart from asking supervising principals for test score data, she has access to district sources of test score information, whereas others reported that applicants occasionally bring reports with their students’ outcome data to interviews.

In addition to reviewing test score data when assessing applicants’ files, principals of higher and lower performing schools alike report looking for evidence that candidates can make data-driven instructional decisions and use student achievement data to improve their own instruction. A principal of the top-performing elementary school in her district said that in selecting teachers, she

wants someone who either knows or is willing to learn to analyze student-achievement data to identify intervention groups and then put interventions in place, someone who knows how to or is willing to learn how to differentiate instruction for above-level and below-level learners.

A principal of a C school asked prospective teachers “how you use data?” during the interview. In reflecting on what she values in candidates’ answers, the principal stated, “The [successful candidate’s] answer would be ‘I like to look at data to tell me how I’m teaching.’ . . . But it also should tell you ‘If I have a child who is not getting it, that child needs differentiated instruction.’”

In making initial teacher placements, school leaders regularly report trying to honor the grade preferences of newly hired teachers as much as possible. But, some principals report using test score data here as well. They “look at the person’s successes. How did their students perform on the FCAT or the writing test? We look at their strengths and weaknesses to decide where we are going to put people.”

Assignment

No school leader interviewed reported feeling constrained in terms of his or her authority to make staffing assignment changes (see Table 2). According to one principal, “As an administrator, you can move anybody anywhere as long as they are certified in that area.” Indeed, the principals in the study estimate their average number of reassignments at somewhat less than 1 to 8 per year, with principals of the lower performing schools reporting higher estimates relative to staff size (see Table 2). Teachers confirm these estimates as well as evidence that principals make assignment changes freely. In making assignment changes, principals report combining teacher preferences with student test score data as they move teachers among grade levels and subject areas.

Grade level assignments. As they reassign teachers to different grades, principals report weighing teacher preferences and evidence about teacher effectiveness, especially student performance. Indeed, all principals in our sample—those in higher and lower performing schools alike—told of using student test score data in their assignment decisions (see Table 2):

In reviewing, at the end of the year, and making plans for the upcoming year, I look at teachers’ test data, their performance. If you have it over time, then a recommendation could be made to see if that person would be better suited [for another grade].

According to an assistant principal in a second district, “[Performance] data is [sic] a big part of it. I would say if we look at something data-wise and we see that it’s not working, a change could come in your assignment, maybe to another grade level or another subject.” The principal of the lower performing school in that district similarly reported, “Some [teachers], because they are not doing well in that grade level and test results show it, I move them.”

In examining her teacher assignments each year, a principal of an A school first asks for teacher preferences and will grant those preferences if FCAT data support them. In her words,

[I ask], “What are your three preferences for teaching next year—first, second, and third?” And, if I feel that it’s a good match for their strengths, I do try to honor [one of] the first two. . . . Now, if a teacher is struggling in that grade level and one indicator [of that] is their [sic] FCAT scores, I have transferred a teacher out.

In a second A school in another district, the principal reported,
Towards the end of the year, I give the teachers a form that says “Where do you want to be next year? What do you want to teach? What grade level? What subject?” . . . Then, I look at my data. I look how the kids did on the FCAT and Dibbles or Stanford Achievement Test. I look at how they did in certain subjects.

As illustrated in Table 2, administrators in 6 of the 10 schools—3 higher performing and 3 lower performing—have formal systems in place each spring through which teachers express their preferred assignments for the following year. According to school leaders and teachers alike, the FCAT plays into teacher preferences as well. At an A school in a large, urban district, the assistant principal recollected that “two requests last year were from teachers who were teaching what we call ‘FCAT grades’ [3, 4, and 5]. They wanted to go from either third, fourth, or fifth grade to K, 1, or 2. I guess there is some kind of, I don’t know, invisible stress of the FCAT.” Interviews with teachers indicate that they distinguish between FCAT and non-FCAT grades and that the distinction shapes their preferences: “You know what, it’s so funny because when I tell other teachers I love third grade, they think something is wrong with me. I like the stress. I love the fact I have to prepare the kids for the FCAT. Not too many people like doing that.”

In some lower performing schools, principals report moving to FCAT tested grades (i.e., 3 through 5) teachers whose previous students made robust achievement gains. The principal of a C school recalled moving eight teachers to new grades or subject areas in the most recent school year: “If I know a teacher is really good, and since third, fourth, and fifth grades are the grades you have the FCAT tests and I really need a stronger person there, I will switch people around.” Last year, after professional assistance failed to improve their students’ performance, two teachers in this school were moved from tested grades to lower ones: “One teacher was in fifth grade and one was in third, and these are crucial grades. We decided that it would be best . . . to have them move down to a lower grade level.” Teachers at the school confirmed the practice:

Last year they did a lot of reassigning. They took a couple of teachers that were in the higher [grade] levels and moved them to the lower levels. The rationale?

You know, those that had good skills could move up to the higher grades and the students would benefit from that, and those that might have been lacking went to the lower grades.

In making reassignment decisions, the principal of a second C school asked herself, “Is this person getting a lot of results in this area where they are? Could they be of help here?” Continuing, she said, “A few years ago as I looked at our school data, our reading was weaker at fifth grade. I said, ‘We’ve got to do something to bring this up!’ I had a very strong fourth-grade teacher who was interested in going to fifth. I made that move.”

By and large, principals in higher performing schools do not systematically move top-scoring teachers to tested grade levels, noting that high-quality teachers are needed at every grade level. According to one principal whose school has moved from a D to an A in the past 5 years, “You can’t say you want your higher achieving teachers in grades three, four, five. If you have high achieving teachers in K, one, and two, then you are going to be okay with three, four . . . . You need strong teachers everywhere.” Indeed, the principals of higher performing schools noted the need to place top-performing teachers in kindergarten through grade 2 where, they said, building reading fundamentals is critical. When asked whether she moved some higher achieving teachers from non-FCAT grades to the fifth grade because of new accountability policies, a principal of an A school replied, “If you don’t teach your children to read in first and second grade, you cannot make that up in third, fourth and fifth grade. . . . So, I have always hired my strongest teachers and put them in that first and second configuration.”

Subject assignments. Other assignment changes are based on teachers’ subject matter expertise. Three of the 10 schools in the study were experimenting with departmentalization as a strategy for improving student performance. When her school’s grade dropped from its previous high the year before, one principal went to a departmentalized structure wherein students in grades 2 through 5 rotate to different teachers for math, science, and reading. A principal of a C school, concerned about her fifth-grade students’ math, science, and reading performance on the state test, also adopted departmentalization:
“I departmentalized [the grade] this year, so I am hoping test scores will be better.” Through departmentalization, principals are seeking to build teacher expertise in particular subject areas, expertise they hope will transfer into higher test scores on state tests in reading, science, and mathematics:

I think that’s what led me back to departmentalized, because some teachers are better at teaching certain subjects. . . . Teachers like to teach everything, but I feel if they are better at teaching one or two particular subjects, I would rather them do that three and four times a day.

Departmentalization in these schools is directly tied to subject matter testing and performance. Because “our science score was low in fourth and fifth grades,” according to the principal whose school grade slipped last year, “I now have one teacher who teaches science four times a day.”

In the three schools that have adopted it, departmentalization has added a level of complexity to the ways in which student performance data are used in teacher assignments. According to an assistant principal involved in those decisions,

Once we research [certifications held by teachers], then we look at data from previous years based on FCAT administration and the SAT-10. We look at how their students in past years performed on different subject areas of the test. We notice if all of their students have done well in math, we would probably consider you to teach math for all of one grade level. So, we have to consider all of that in assigning teachers and assigning by subject area.

In a fourth school, the principal, citing pressure from accountability mandates, has begun to implement small-scale, voluntary departmentalization on a classroom-by-classroom basis when teachers are supportive. When asked how he assigns teachers, the principal, fighting to retain his A status, said,

Well, it depends. No Child Left Behind—we’ve had two teachers who agreed to be co-teachers in fourth grade who like to swap off reading and math. So, two teachers: one might teach reading to both groups of kids, and one might teach math. . . . We try new things every year, and tweak it constantly. Our scores are way up, and we thought some of the things we did last year may have been responsible.

Departmentalization in some districts has become ubiquitous and, in some schools, has been implemented with the youngest of children:

All of the principals I talk to . . . are doing departmentalization. Some are doing it kindergarten through fifth grade. For me, I thought it was too much movement for K and one, so I have them self-contained. We made the decision to do it second through fifth this year. Of course, I will look at my scores—my Stanford [SAT-10] and my FCAT for second and third [grades]—to see if that worked.

Teachers in these schools confirm both the departmentalizing practices and that student performance on state tests had led to the change. They also describe a mixed reaction from teachers:

Teacher: At first, we didn’t want to [departmentalize]. It was like the administration didn’t give us a choice: “Well, we are going to do it,” but the teachers were very unhappy about that at first. Very.
Interviewer: Now that it’s implemented, do you think there is a general sense that it works?
Teacher: We will see when the test scores come.
Interviewer: Do you think that test scores drove that structure change?
Teacher: I know it [sic] drove that change.

Interviews with teachers further indicate that departmentalization in two of the schools has been coupled with tracking for some tested subjects. From a teacher in a second school,

Teacher: We are grouped this year.
Interviewer: In terms of achievement?
Teacher: Yes. FCAT scores. My homeroom class, I have the lowest. But, we are also departmentalized so even if you have the lowest in the morning, in the afternoon you will get a top class to teach the same thing.

At an F school, where student math scores were contributing to a low school grade, the principal had just implemented class-by-class tracking for mathematics instruction in third grade as part of a new supplementary mathematics curriculum.

Development

All principals interviewed report regularly using student FCAT scores or other assessment data to develop individual teachers in their schools (see Table 2). Specifically, they report using data from the state test and benchmark assessments to
plan classroom observations and focus professional development opportunities; with the exception of school leaders in one district, principals deny that they have transferred teachers out of their schools for performance reasons.

Student outcome data help to guide principals’ decisions about evaluation schedules and whose classrooms to visit:

We do provide a lot of support to the teachers during their first couple years here. If you do not succeed, it’s not because there has not been a lot of support given to you. . . . I measure success by going in and out of their classroom and seeing if their instruction is on target. . . . And, test scores. We do K-3 testing with Dibbles, and so I can see changes from the first year. . . . And certainly, the FCAT with fourth and fifth grade, your children should be making gains.

Principals also report using other data in “real time” to work to improve instruction throughout the school year. The principal of an F school says she uses student test scores on curriculum-embedded assessment modules to identify training needs. She recounts how the assessments told her, for example, that students were not yet competent with a writing standard that would be assessed on the fourth-grade FCAT. In response, she provided training on the standard to all fourth-grade teachers.

Principals and teachers alike say that FCAT scores guide professional development decisions. In developing proposals for annual individual professional development plans, for example, teachers are required to demonstrate how planned training aligns with their instructional deficits:

Our county is very adamant about teachers constantly looking at last year’s [student test] scores and adjusting their instruction accordingly. Teachers have to look at last year’s students and decide how they could teach this year’s students better to meet some of last year’s deficiencies.

With the exception of one, all principals interviewed denied transferring teachers out of their schools for performance reasons (see Table 2); instead, they reported investing in them. According to one principal whose students’ passage rate on the FCAT in reading and mathematics has moved from 45% to 90%, “My perception is you develop your people, and either they work and grow, or they . . . get on a professional development plan and make those improvements or you go through the [termination] process.” Continuing, she said, “To me, if an employee has issues or difficulties, you [the principal] need to be responsible for working with them. I don’t even see an option being: transfer them somewhere else and let someone else deal with it.” When asked whether he had transferred teachers out of his school for performance reasons, a second principal whose school had improved from a D to an A replied, “I’ve never considered it. If I can’t fix a teacher myself, I’m not a good principal.” Interviews with teachers across the seniority spectrum corroborate principal accounts.

The only exception to this practice came from principals and assistant principals interviewed in one of the state’s lowest performing districts. In responding to questions about involuntary teacher transfers, principals in the district referred to the pressures placed on them by the state’s grading system: “The biggest thing in the State of Florida now is the FCAT, and if a particular teacher’s FCAT scores is [sic] plundering, then you know you got a problem there. . . . You know some schools would unload that problem [onto another school].” In response to a follow-up question about whether the accepted practice in the district is to dismiss untenured teachers who for performance reasons are not recommended for annual contract renewal or to place them in other schools in the district, the principal replied, “I have seen it work both ways.”

The sole time a principal in the second school did not recommend an untenured teacher back for the following year, the superintendent transferred the teacher to another school in the district: “I am kind of embarrassed for us that he actually was hired back, and at a double F school!” In explaining the decision, she continued,

We are in a situation, district-wide, where you may see some of the things that you wouldn’t even hear of in other districts. It’s because of the [teacher] pool and we are limited in candidates. . . . Sometimes they are extremely qualified, and sometimes it’s the ones that are left over from other counties who have not received a job yet.

An assistant superintendent in the district echoed this sentiment. When asked about the pool of teacher candidates, she said, “The criteria for assignment is [sic] certification and a pulse!” At the school, principals and teachers report that poorly performing tenured teachers are also transferred in lieu of being terminated. With 4 years experience at the school, the assistant principal reported that one tenured teacher, on
average, is transferred out of the school for poor “performance on test scores and job performance in general” per year.

Dismissals

Student test score data appear to play a more limited role in teacher dismissals. According to district and school leaders in all five school districts, tenured teachers are infrequently dismissed from their positions for performance reasons. In fact, no principal interviewed reported ever dismissing a tenured teacher for poor performance (see Table 2). Site union representatives in the schools confirmed these reports. A union representative who taught in the same school for 28 years, for example, said, “I don’t know of any continuing contracted person that’s ever been terminated because of performance. I don’t know of any.”

In only one district do administrators attribute this to teacher supply. Those in the other four districts give two other reasons. First, they say, are stop gap measures that assure poor performers are weeded out early and do not become tenured. In Florida, stop gap measures include the “97 day rule” adopted by the state legislature in 1997. It is a probationary period during which first year teachers may be fired without reason. Principals admit that their hiring methods are not fail-proof, and the probationary policy is easy—“There’s no documentation tied to that 97 days”—and allows them to correct their mistakes:

You can do an interview with a person, but you cannot dictate how that particular person is going to fit within the classroom itself because you have people that are great interviewees that don’t perform well. . . . The State of Florida has what they call a 90 [sic] day window for a beginning teacher. And, it says that within those 90 days, you can release a teacher, and you don’t have to give them a reason.

A second principal, who released two teachers under the 97 day rule in her 3 years as principal, admits that such decisions are made, to a large extent, in the absence of performance data: “I have released teachers within those first 97 days, and you don’t have a lot of data. Last year, I let two go within that time frame. But, you know, you look around and see kids are not learning, and you have to make some changes.”

In addition to the 97 day rule, principals also point to their dismissal options during the pre-tenure period. At a turnaround school, the principal said simply, “If they’ve not worked out at the end of any of those three years, you can simply let them go with no documentation. After three years you should be pretty sure that that’s who you want to keep.” She reported that during her 6 years as principal at the school, she has not renewed contracts for, on average, one or two untenured teachers per year. According to a principal of a lower performing school in another district who reported not renewing contracts of, on average, one untenured teacher annually, “At the end of a yearly contract, there’s simply a letter that your annual contract is not renewed, and your contract will end on such and such a date.”

Interviews with principals suggest that decisions about whose contract will not be renewed next year are informed by performance data. When asked on what bases nonrenewal decisions get made, an assistant principal in an F school where, on average, four to five teachers are not asked back per year stated, “Primarily teacher data, but sometimes school need, school fit, classroom management, if that was a problem. Those are the main factors.” Asked the same question, a principal of another lower performing school replied, “It was overall student performance, either looking at test scores and report cards or their daily performance.”

Second, the principals report that dismissing tenured teachers is unnecessary because of resignations by teachers placed on professional improvement plans (also called success plans). Principals say that the process leading up to dismissal is so rigorous and the stigma associated with dismissal so severe that poor performing, tenured teachers simply resign before they can be dismissed. Of the three tenured teachers a principal at a turnaround school placed on professional development plans, all three resigned:

There’s a process. You complete professional development plans with them, and it’s all spelled out. If teachers complete the professional development plan and it’s satisfactory, then we move on. If not, then dismissal becomes a recommendation to the superintendent and the school board. We’ve never gone past the professional development plan. The three that I’ve had [on professional development plans] opted to resign.

497
Discussion

Overall, findings suggest that staffing, in the context of state accountability pressures, is at least in part “evidence based.” It is evidence based in the sense that principals report using students’ scores on the state standardized exam and alternative tests in non-FCAT grades (e.g., Dibbles; SAT-10) to make teacher staffing decisions. If we are to believe principal reports, school leaders in Florida today do not hire on “instinct” and “good fit” alone. Nor do they simply give senior teachers assignments of their choosing without consideration for the effect on test scores. Professional development, too, is planned with student assessment data in hand, even as it appears that the termination of tenured teachers remains somewhat recalcitrant to accountability pressures.

Of course, knowing what principals say they do is not the same as knowing what they actually do. In this sense, the article represents a good starting point for further work to probe the role of student performance data in staffing behaviors. For example, future research might calculate the rate of teacher reassignments in low- versus high-performing schools or longitudinally in individual schools where performance (e.g., school grade) has changed over time. Within schools, studies might consider which grade levels see the most switching in order to examine whether principals’ reported focus on tested grades and subjects holds up in practice. Other work might combine teacher level test score data with school personnel documents or administrative data to explore whether higher performing teachers are reassigned to grades and subjects that pose specific achievement challenges in the schools where they teach.

Only one study has looked at the relationship between student achievement and observed teacher reassignments. Chingos and West (2011) followed the career paths of nearly 25,000 Florida teachers for 7 years beginning in 2001–2002. They found that effective teachers, as measured by value-added estimates that regress students’ math and reading test scores on their prior-year scores, are less likely than ineffective teachers to be reassigned to teaching positions in untested grades or subjects. Moreover, they reported that teachers in the top value-added quartile are most likely to remain in high-stakes grades and subjects in schools that face strong pressures to improve performance, as measured by the ratings they receive from the state’s accountability system.

Test-based staffing practices documented here and by Chingos and West (2011) appear to represent a significant departure from the conflict-avoidance and seniority-based approaches that the admittedly limited literature on the topic suggests characterized personnel decisions in the past. They might be viewed as another example in a growing list of ways schools, under unprecedented pressure for achievement and with relatively little empirical evidence about what works, have become incubators of new practices. According to Crowson and Goldring (2009), a new localism in educational governance recognizes that the growing national and state prescriptiveness in educational objectives is in fact pulling school actors ever more fully into the particulars of their own local contexts. In their words,

while state and federal mandates in education are strengthened centrally, the nation’s prime attention under these mandates is determinedly local. School districts, and especially individual schools, are the focus of accountability pressures, of improving-achievement, and the center of a renewed search for both teacher and administrator effectiveness in instruction. (p. 2)

Cohen-Vogel and Rutledge (2009) show that this new kind of localism situates instructional arrangements that support student achievement at the center of efforts for school change. According to the authors,

Content coverage, instructional time, even pedagogy all show signs of “give” in the context of recent accountability pressures. No longer do teachers simply “shut the door” and teach “what” and “how” they will. Nor is instructional time divided up more or less equally among subject areas. (p. 87)

Findings here suggest that staffing practices, at least in Florida, also appear to be responding to the pressures.

Even as staffing decisions may be becoming evidence based, they do not yet reflect prescriptions for individual teacher–student matches. Although principals in Florida report that they are moving...
teachers to different grade levels and subject areas on the basis of student test score data, there is little evidence to suggest that they are using the same data to reassign teachers to particular students. Put another way, there is not yet an equivalent in staffing for practices and technologies, like Assessment to Instruction (A2i) and others described by Connor, Morrison, Fishman, Schatschneider, and Underwood (2007), that prescribe select curricular activities (say, reading comprehension practice) for each individual child. Participants in two schools with new departmentalized structures report that school leaders there use performance data to group students into homogeneous classrooms, a finding that did not reflect the experiences of educators in the other elementary schools. In the remaining eight schools, as my colleague and I show elsewhere, test score data are used to inform principals’ decisions as they assign students to classes, but they are used to ensure heterogeneity and a sense of “fairness.” In fact, principals use index cards with information on student gender, race/ethnicity, and achievement (i.e., FCAT scores, student grades) to populate classrooms that spread high, middle, and low performers equally across teachers (Osborne-Lampkin & Cohen-Vogel, working paper). The practice of spreading student ability equally across teachers is pervasive in the sampled schools and assumes that “good”/“bad” teachers are good/bad for all students equally. That is, it suggests that principals believe that teachers are equally effective with different types of students; there is surprisingly little research in the area of differential teacher effects to tell us whether they are right.

Qualitative approaches are not ideal for measuring the prevalence of a phenomenon but are instead most valuable for theory development and to build understanding about how a phenomenon—in this case, teacher staffing—works. Although my data, drawn primarily from semi-structured interviews, reveal important new insights into the ways principals may be using achievement data in their staffing decisions, they do not provide strong indicators of the extent to which principals use data in these ways. Future research employing survey and related methodologies is needed to gauge just how widespread test-based staffing practices have become. Surveys may also be able to address whether evidence-based assignment decisions of the type documented here are limited to elementary schools where general elementary certification makes teachers easier to move around and whether they are unique to Florida and, perhaps, other high accountability states where school recognition funds are distributed to schools based on school grades and teacher-level student outcome data are readily available.

More work will also be required in order to substantiate whether and how staffing strategies vary between higher and lower performing schools. Findings here suggest that, whereas principals across school performance levels appear to be using test data in their staffing decisions, differences may also exist. Principals of lower performing schools and schools whose performance has dropped appear more likely than others to departmentalize instruction and move effective teachers to tested grades, for example. Why? One explanation is that principals in these schools may feel less certain about their job stability and, as a result, be more willing to innovate and face the political “heat” that moving teachers and departmentalizing instruction can generate in an effort to turnaround test scores.

If staffing differences between higher and lower performing schools are substantiated in future work, what are some possible implications? Like others, Ellis (2008) pointed to studies that show it is sanctioned schools—schools that also educate a disproportionate number of poor, minority students—that are more likely to direct instructional time to test preparation and reading instruction (Anagnostopoulos & Rutledge, 2007; DeBray, Parson, & Avila, 2003; Diamond & Spillane, 2004; Jennings & Stark, 2006; Lemons, Luschei, & Siskin, 2003; Rutledge, 2008; Tate, 2000). He worries that current efforts to raise test scores may, in his words, be academically restraining because they are built on a model of learning that discounts learners’ understanding of [the subject] in order to privilege their relative standing within an artificially “standardized” set of boundaries, thereby limiting actual opportunities for student success in higher level coursework (Tate, 1994, 2000). (p. 1342)

With high-poverty, high-minority students more likely to attend elementary schools facing pressures to improve, it is their classrooms that
may be becoming departmentalized. Departmentalization, according to some, provides kids with fewer opportunities to make cross-disciplinary connections. Moreover, the tracking that, at times, accompanies departmentalization has been shown to lead to powerful within-grade differences in the time students spend reading and writing, their access to challenging material, and the quality of teacher feedback (Lucas, 1999; Oakes, 2008; Watanabe, 2008). Furthermore, if our early findings hold, it appears that it is in high-poverty, high-minority schools that the most effective teachers are being moved out of kindergarten through second grade, where children are expected to master the fundamentals of reading that, if not well-obtained, threaten long-term academic success (see, for example, National Research Council, 1998). From this perspective, differences in the ways higher and lower performing schools may be allocating instructional resources in an effort to improve student achievement raise new questions about the extent to which No Child Left Behind and the standards and accountability frameworks that continue to dot the state and federal policy landscapes address deeply rooted structural inequities of race and class in America’s public school system.

Acknowledgements

The author gratefully acknowledges research help from La’Tara Osborne-Lampkin. The statements made and views expressed are solely the responsibility of the author.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported in this article was made possible by a grant from the Spencer Foundation. The views expressed are those of the authors and do not necessarily reflect the views of the Spencer Foundation.

Notes

1. A rural district in the Florida panhandle with approximately 180 teachers, the Calhoun School District, has no organized teachers union.
2. Charter schools were eliminated as they are often exempt from policy and procedures (e.g., collective bargaining provisions) that could affect the ways in which principals staff their schools.

References


Goertz, Goe, Garet, 502
Fetterman, Figlio, Cohen-

Staffing to the Test


503


Tucker, P. D. (1997). Lake Wobegon: Where all teachers are competent (or, have we come to terms with the problem of incompetent teachers?). *Journal of Personnel Evaluation in Education*, 11, 103–126.


This content downloaded from 149.175.1.33 on Mon, 12 Jan 2015 16:05:45 PM
All use subject to JSTOR Terms and Conditions
Staffing to the Test

tracks in North Carolina. Teachers College Record, 110(3), 489–534.

Author

LORA COHEN-VOGEL is the Robena and Walter E. Hussman, Jr. Distinguished Professor of Policy and Education Reform at the School of Education, University of North Carolina at Chapel Hill. Her research focuses on the politics of PreK-20 reforms, and the programs, policies and practices that characterize schools that are successfully raising outcomes for traditionally underperforming students.

Manuscript received July 1, 2010
Revision received January 8, 2011
Accepted April 1, 2011