

# Chapter 2

## The Press From Above, the Pull From Below: High School Responses to External Accountability

Diane Massell, Margaret Goertz, Gayle Christensen, and Matthew Goldwasser

### Introduction

In the research literature, high schools are often portrayed as the level of the educational system most resistant to reform initiatives. Unlike elementary schools, for example, high schools are balkanized into subject matter departments, teams, academies, and other substructures, making communication and influence very complex and challenging (see Siskin, 2004). Perhaps because of their reputation for intransigence, high schools have received comparatively less scrutiny and focused attention from policymakers than elementary and middle schools.

Recent state standards-based reform initiatives, however, do include high schools under the umbrella of performance accountability. In contrast to conventional efforts to monitor school compliance with input and process regulations, these newer forms of accountability focus on student academic outcomes, schools' continuous improvement on explicit performance targets, the public reporting of test results, and greater consequences for failure to succeed (see Fuhrman, 1999). Many states have also attached incentives for high school students to improve their performance on state tests, ranging from scholarships to grade promotion or graduation.

This chapter explores high school teachers' and school and district administrators' response to their state accountability system. This system includes not just the particular design of

accountability programs, but also the state tests and content standards to which they are tied. In the following section, we offer an overview of those systems for high schools during the time of our fieldwork in 2002–2003, the year preceding the implementation of the No Child Left Behind Act (NCLB). We evaluate the strength and stability of state accountability programs, and compare the nature of different state assessments and content standards.

Then we look at whether educators were aware of and understood the accountability expectations, and explore their perceptions about its value. We anticipated that educators' knowledge and awareness of the system and its demands (Abelmann & Elmore, 1999; Elmore, Abelmann, & Fuhrman, 1996) would contribute to a greater likelihood that they would act on its behalf. And ever since the early RAND studies of program implementation in the 1970s, it has become a truism that action on behalf of policies and programs is highly dependent upon the will and commitment of the “street-level bureaucrats” expected to carry them out (Berman & McLaughlin, 1975; Lipsky, 1980). As expressed by one of the original authors of the RAND reports 15 years later:

*Policy cannot mandate what matters [italics added]. . . . The presence of will or motivation to embrace policy objectives or strategies is essential in the generation of the effort and energy necessary for a successful project.*  
(McLaughlin, 1990, pp. 12–13)

This view is now deeply embedded in reform strategy. So, for example, when national and state education policymakers decided that content standards were an essential tool for school improvement, they tried to gain the consensus of the public and professionals through an inclusive standards development process. They assumed that such participation would produce the buy-in needed for strong implementation (Massell 1994, 2000).

Next, we discuss the salience of the accountability system to high school educators and administrators, e.g., the extent to which high school educators and administrators reported feeling pressure from the accountability system, and acted in response to its demands. The salience of the system predictably differed across schools. At the low end of the spectrum were schools where individual actors felt some pressure and took some initiative to address accountability demands, but the departments or school did not act as a whole. By contrast were schools that collectively launched one or sometimes many accountability-related initiatives.

In the remainder of the chapter we analyze how these various factors—accountability system designs, knowledge and perceptions, and salience—and others that emerged as important (notably, district press and school leadership), contributed to high schools' response to their state's accountability system.

## **Standards, Testing, and Accountability for High Schools in the Six States**

Accountability policies, as applied to high schools in 2002–2003 in the six states selected for study, varied on a number of dimensions that could affect educators' understanding and acceptance of, and response to, state policies. These include the target of the accountability system (student, school, and/or district), the type of assessments used for student and institutional accountability, the nature of the accountability measure and its consequences, and the overall strength and stability of the accountability system. Table 1 summarizes these variables across the six states, and Table 2 provides specifics about each state assessment; the tables appear at the end of the chapter. Note that we gathered information on state assessment and accountability policies from extant reports (Center on Education Policy, 2003; Goertz & Duffy, 2001), state Web sites and published information, and follow-up interviews with personnel from state departments of education. As described in greater detail later, these policies represent variation in the incentives used to capture schools' attention, the assessments used to measure student performance, the support provided for school improvement, and the historical and political contexts of the policies.

In this section, we look across state policies to discuss the strength of the accountability systems (as measured by the target and consequences of the policies) and their stability, the nature of the state testing program and its perceived alignment with state standards, and the specificity of state guidance.

## Strength and Stability of Accountability Systems

We categorized the six state accountability systems by the scope of their coverage (student, school, district) and the nature and strength of the consequences applied to the accountability targets. We consider four of the states—California, Florida, New York, and North Carolina—to have strong accountability systems, and the other two states—Michigan and Pennsylvania—to have weak accountability systems.

### Strong Systems

The four strong systems each held high school students accountable through a high school exit examination. This requirement had been in place in Florida, New York, and North Carolina since at least the mid-1980s, although all three states have increased the rigor of their high school assessment over that period of time. Passing the Florida Comprehensive Assessment Test (FCAT) became a requirement for class of 2003, while passing the Regents Comprehensive Exams (RCEs) applied first to the class of 2000, and the Competency Test in North Carolina went into effect in 1994. These assessments replaced minimum competency tests as high school graduation requirements in all three states. In addition, North Carolina high school students take a series of End-of-Course (EOC) exams that count for 25% of their course grade. California's High School Exit Examination (CAHSEE)—the state's first—was implemented in 2000–2001, and applied to the class of 2004 at the time of our fieldwork. In July 2003, just after the completion of our site visits, the State Board of Education voted to make the class of

2006 the first students subject to the graduation requirement.

These four states also held high schools accountable for student performance through a combination of rewards for high and/or improved performance and sanctions for low performance. Although the formulas for identifying schools differed across the states, all four took into account both the absolute level of student performance and changes in achievement over time. Low-performing schools received technical assistance (TA) through state assistance teams (NC, NY) and additional funds (CA, NY). Schools that failed to improve over a designated period of time could lose their students (FL), principal (NC), or accreditation (NY), and/or be subject to reconstitution (FL) or state takeover (CA). While the school was the primary target of accountability in these states, districts in North Carolina could have their superintendents and other administrators replaced and lose their accreditation if over half of their schools fell into the lowest performance category.

Accountability policy in California, Florida, New York, and North Carolina was relatively stable. High school students in the last three states had been subject to a high school exit examination for many years. Florida and North Carolina's school accountability systems were also at least a decade old, and could be characterized as "mature." North Carolina's ABCs program was enacted in 1994. Although Florida's A-Plus program dates only from 1999, it refined and expanded an earlier accountability policy that focused on the state's lowest performing schools. As California and New York's school accountability policies are of more recent vintage, we have designated them as "emerging." California's Public School

Accountability Act was passed in 1999. New York has held its very lowest performing schools accountable for several years under its Schools Under Registration Review program, but only extended its accountability policies to all schools in the state in 2000.

## **Weak Systems**

While high school students in Michigan and Pennsylvania must also take state assessments, passage was not a state requirement for graduation. Students who passed the Michigan Educational Assessment Program (MEAP; Basic or above) received an endorsement on their diploma and could qualify for a college scholarship (score at Proficient or above). In Pennsylvania, local school districts determined whether to use the 11<sup>th</sup>-grade (Pennsylvania System of School Assessment (PSSA) and/or a local assessment as a graduation requirement. The state had considered, but not implemented, a policy of rewarding students who passed the PSSA with diploma seals.

At the time of our fieldwork, Michigan did not hold high schools accountable for student performance. The state's accountability system was initiated in 1994, but was placed on hold in 2000 while the state designed a legislatively mandated system based on multiple indicators of student performance and school context variables. State accountability ratings were not released until fall 2004. Michigan did not have a system of district accountability, either. Pennsylvania, however, held districts accountable for aggregate student performance. Under the state's Empowerment Act of 2000, districts in which half of the tested students score Below Basic on the state assessment were

subject to a series of sanctions, from the development of a district improvement plan to technical assistance to state takeover. Schools could receive rewards for improved test performance and/or attendance, but were not subject to sanctions.

## **State Assessments**

High school assessments in the six study states varied on several dimensions: the number of assessments, grade level tested, content coverage, and remediation policy. All six states asserted that their tests are aligned with state standards.

Three states—California, New York, and North Carolina—administer multiple tests to high school students that are, in turn, used for different accountability purposes. High school students, for example, take the CAHSEE in grade 10, the California standards tests (CSTs) in grades 9–11, and the California Achievement Test (CAT-6) in grades 9–11. While only the CAHSEE counts for high school graduation, all three tests are used to calculate a high school's accountability index. Similarly, in North Carolina, only the eighth-grade End-of-Grade (EOG) exam is used for high school graduation. Scores on end-of-course (EOC) examinations count toward students' course grades, and both of these exams, as well as the 10<sup>th</sup>-grade High School Comprehensive Test, are used to calculate a school's status and growth rate under the ABCs. The New York RCEs are given in multiple subjects. Students in the class of 2001 had to pass two exams (English and mathematics), and, starting with the classes of 2003 and 2004, students must pass five examinations (two additional in social studies and one in science). The other three states administer only one high school examination.

The high school exit exams are administered at different grade levels: grade 8 (NC), grade 10 (CA, FL), and at the end of the specified course (NY). Similarly, EOCs in North Carolina are administered when students complete a tested course. Michigan and Pennsylvania administer 11<sup>th</sup>-grade high school tests. While the high school exit exams tend to be limited to English/language arts and mathematics (CA, FL, NC), high school students are also assessed in science and/or social studies in exams that count toward graduation and/or school accountability in California, New York and North Carolina. While not a high-stakes test at the time of our fieldwork, Michigan's 11<sup>th</sup>-grade MEAP also covers science and social studies.

The grade level and focus of the test (specific course versus more general skills and knowledge) also affect test content. For example, North Carolina's exit exam covers the state's eighth-grade standards, and students are expected to perform at grade level on that material to graduate from high school. California's HSEE is aligned with 9<sup>th</sup> and 10<sup>th</sup> grade English/language arts (ELA) standards and with sixth- through eighth-grade mathematics standards (which include algebra). Michigan's 11<sup>th</sup>-grade tests cover algebra and some geometry, while Florida's 10<sup>th</sup>-grade and Pennsylvania's 11<sup>th</sup>-grade mathematics tests include material through precalculus. North Carolina's EOC exams assess the state's Standard Course of Study in 11 courses in English (through English II), mathematics (through Algebra II), science (through Chemistry), and social studies.

Finally, the four states with high school exit exams require that students who fail the tests receive remediation. Local districts design these remedial programs, which often receive additional funding from the state. Neither Michigan nor Pennsylvania fund nor

require intervention services for students who fail their high school assessments.

## **Standards: Specificity of State Guidance**

In 2002, the specificity of state guidance for the high school curriculum varied as well. At one end of the spectrum is North Carolina, with its course-specific guides and EOC examination structure. The latter provides teachers with specific feedback regarding student performance on a bounded content area, and gives school and district staff detailed information to help them adjust instruction to meet state standards. Florida's curriculum frameworks, curricular planning tools, and course descriptions incorporate grade-by-grade state standards. California has high school standards that focus on grade clusters (9<sup>th</sup>–10<sup>th</sup> and 11<sup>th</sup>–12<sup>th</sup>), but more specific teacher guides to assist in preparation for the state high school exit exam. The latter generates subject-area scale scores and subscores on the skills and content of that test. By contrast to these states, the guidance offered by Michigan and Pennsylvania is much less specific. For example, both states' standards documents and curriculum frameworks were only for benchmark grades, not for each high school course or grade level.

## **Knowledge of State Accountability and Its Perceived Value**

The preceding discussion reveals just how complex and multifaceted some state accountability systems can be. Nevertheless, a majority of high school teachers across the states could paint at least a broad portrait of state accountability. Most had a clear

understanding of the potential consequences (or lack of consequences) in the system, and the elements for which they were directly responsible, such as student testing or documentation of curriculum alignment. Misconceptions or confusions about accountability were more common among educators in Michigan and Pennsylvania, where the accountability policies were in flux. For instance, Pennsylvania teachers were confused about the status of diploma seals, which the state had rescinded after a year, and did not understand the sanctions and rewards. Notably, Pennsylvania, unlike other states, applied accountability consequences primarily to districts rather than schools.

Predictably, staff roles and responsibilities mattered in terms of the specificity and depth of their knowledge about accountability design. Principals, school improvement team members, and in some cases department chairs had a more complete and detailed picture of the system. District officials understood its complexities well. Of course, school and district administrators are usually responsible for submitting accountability data and reports to the state, and must answer to local school boards about schools' progress, so it is not surprising that they have more intimate knowledge of accountability.

As noted earlier, we were also interested in our educators' opinions about their state accountability system, anticipating that these views would shape the extent to which they responded. In what follows, we discuss their views of each different component of accountability.

## Standards

Educators' views about their state content standards differed across the states. In California, North Carolina, and

Pennsylvania, at least,<sup>1</sup> content standards appeared more highly regarded than the states' testing or accountability programs. For example, the majority of North Carolina teachers believed that state standards accurately reflected what students should know and be able to do.<sup>2</sup> Many said that the state's Standard Course of Study was the foundation for their teaching, and one referred to it as her "Bible." Teachers even in nontested subjects used the state's Standard Courses of Study as a guide for their instruction.

In addition, teachers argued that content standards produced a more coherent curriculum and more consistent expectations. Said one California teacher:

*No, I really don't, I really don't think it's [standards are] a negative. I can look at the standards and I can say, yeah, kids should know that. And we just can't go through education with a hit-and-miss thing, where one kid gets into a particular teacher's class and learns a lot, [inaudible] another class and learns nothing except for what that teacher did over the weekend. And it's got to be more, if we're going to experience gains, it has to be more than just a hit-and-miss thing. (Teacher, San Antonio High School, CA)*

At least rhetorically, standards were often viewed as more legitimate goals for teaching than tests. One New York department chair, echoing a common sentiment in the school, stated emphatically that "lessons should be standards based. We're not teaching to the test.

---

<sup>1</sup> Educators in the other states did not express much opinion one way or another about their state standards.

<sup>2</sup> There was one exception to this sentiment; Maple High School teachers did not think the standards were not adequate.

We're teaching to the standards" (English Department Chair, Nelson High School).

Interestingly, however, this school did conduct quite a lot of test prep activities. We found this schism between rhetoric about testing's value and action elsewhere, across schools with relatively weak and strong performance under their accountability systems.

State standards did not go without critique, of course. In Michigan, state standards for benchmarked grades were seen as too vague to provide useful guidance. (Indeed, the state, regional Intermediate School Districts, and school districts themselves were making an effort to articulate and specify the standards for classroom teachers.) Criticism of content standards was common in Pennsylvania, even though standards fared better than other components of the state accountability system. Pennsylvania has arguably had a more difficult political history of standards development than other states in this sample. For example, when the state first introduced standards in the late 1980s and early 1990s under the name Outcomes-Based Education (OBE), it galvanized opposition from all ends of the political spectrum. The Christian right community rallied against values statements. They were joined in opposition by more liberal groups who were concerned about the state asserting authority over curriculum, and bitter debates ensued. So, for example, staff in Orthodox High School and its district administrators thought the standards tainted by politics, and preferred national content standards. Support for standards across Pennsylvania schools is best characterized as moderate.

Consequently, relatively few Pennsylvania respondents identified meeting standards in their goal statements, or in what they felt accountable for. Such results stand

in sharp contrast to California, where teachers in six of the eight schools felt accountable for covering state standards. Indeed, in most other states, teachers typically said they felt accountable for meeting academic goals and/or aligning their curriculum to state standards, along with creating a love of learning and helping students to achieve their full potential. While improving student motivation, student behavior, and persistence in school were mentioned, they were not as prominent as the academic goals.

It is interesting to recall that when contemporary standards were introduced 10 or more years ago, they often met with extreme resistance even in states with a well-established history of curricular guidance. For example, there was a firestorm in response to New York's social studies standards in the late 1980s (see Massell, 2000). In other states like Colorado, educators and the public were quite wary about the extension of state control into the curricular prerogatives of local districts and schools. But while the road to state leadership in curriculum has sometimes been tumultuous, standards have become an accepted and legitimate feature of state policy in all of our states, with Pennsylvania as a slight exception.

## Testing

State assessments garnered a substantial amount of comment, both positive and negative. Some educators believed that testing set clear goals for students, and provided them with a useful way to calibrate their expectations of student performance and evaluate the success of their own teaching (e.g., Southern High School, NC) or of their departments as a whole (Upstate High School, NY). Said one New York teacher:

*I think it [the Regents exam] really puts an end to social promotion. You know and this is no knock against any teacher but you know sometimes we might be inclined to pass students who have been working real hard and really don't understand the material. And now it's basically saying you can't move on unless you show proficiency on an exam. (Teacher, Nelson High School, NY)*

While she believed the Regents exam compelled teachers to hold back failing students (buttressing the argument that standards set high expectations), another teacher admitted he was more likely to pass a student failing his coursework if the student had an acceptable Regents score. In either case, teachers do seem to agree that standards build greater consistency from teacher to teacher and school to school, as standards reformers have argued (U.S. Department of Education, 1994).

But while educators acknowledged positive aspects of state testing, more often than not they challenged the quality of the tests or their impact on the school's curriculum. Teachers and administrators across the sample states questioned the accuracy of state tests as a measure of student achievement, either because the reliability and validity of the tests were flawed, or because they thought that once-a-year tests did not adequately capture what students knew. The largest area of concern was the impact of testing on the content of the high school curriculum.

New York teachers and administrators were extremely negative about the Regents examinations, particularly in mathematics. They argued that the validity of these exams had declined and that the assessments had become more of an obstacle to be overcome than a measure of true learning (e.g., River City High School). One argued that the math exam was so heavily dependent upon reading that even brilliant math students could perform

poorly (Teacher, Nelson High School). Others complained that the Regents format was mysterious and continuously changing, with sections weighted differently from year to year. Indeed, just after data collection, the state pulled the easier of the two math exams (Math A) because only 37% of students passed. The state subsequently allowed students to graduate and receive local diplomas if their teachers attested that they had met state standards.

Some math educators in North Carolina also challenged the validity of some state tests, but in this case they thought student scores on the state EOC examinations were higher than they should be.

*Those tests aren't real indicative of what kids can do. If I have a student that's getting Fs and Ds in my class, and they can come out with a B on that exam, I have a problem with the validity of that test—I'm not sure how they score them, you know? And I have had that happen lots—kids who have gotten Ds and Fs because they do absolutely nothing, and I can see them getting a D on the end of course exam, but then come up with a B. I've even had kids come up with an A on it. (Teacher, Lincoln High School, NC)*

Pennsylvania teachers thought that neither very low nor high end students were well-served by PSSA. Some Michigan educators observed that their scores had fluctuated widely from year to year. They thought these shifts were due to unreliable scoring and scaling practices rather than changes in their instruction or their student population. They and others noted that state tests were administered too infrequently, with results returned too late, to be of much use for improving classroom practice, countering the oft-stated policy expectation that test data would be a critical lever for instructional improvement.

Many high school educators across the states thought that state assessments had a negative impact on the content of their

curriculum, and observed a variety of narrowing effects. Nontested subjects, such as foreign languages, reportedly received fewer resources under the new accountability regime. Teachers in tested subjects argued that they had had to reduce their curriculum to topics covered by the state test; some felt that the state tests forced them to cover too much, too quickly, at the expense of diminished student learning and dampened instructional creativity. Another kind of narrowing was experienced by students in some of our Florida and California schools; these students were reportedly required to take the same or similar courses over and over until they passed state tests. An assistant superintendent in one of our Pennsylvania districts captured the spirit of these concerns when he said that the state's focus on testing led to a "teach-to-the-test-at-the exclusion-of-the-meat-of-learning program" (Orthodox High School, PA).

## Accountability

Of course, state tests were the primary performance indicators used in state accountability programs, so some of the positive and negative comments about the tests had close parallels with educators' views of the accountability measures and consequences tied to the tests. Administrators and teachers expressed the sentiment that the *idea* of holding high schools accountable for test results was a legitimate expectation and could be useful in motivating them to reflect upon the effectiveness of their practices for student learning. For instance, one Florida principal said:

*But we have to start someplace, and not to start would be more detrimental. So I feel for some of the schools who are really having a difficult time with this test area, the grading and all that. But I also feel like we have to have something in place to make*

*sure that we provide what's best for the kids.*  
(Principal, Harbor High School, FL)

Similarly, a Michigan principal said his staff would become complacent without accountability or their regional accreditation process (Principal, Smith High School, MI). Said one California school administrator:

*Are they progressing? Are they progressing fast enough? Are you presenting a rigorous curriculum? I mean, that's another word that we were using a lot last year, and I like it. And, sure, you could be a real nice teacher and very nurturing, but are you rigorous enough? Are you having them write too much in journals, in their personal...you know, I think journal writing is good, but are you have them write expository essays? Are you providing enough challenging reading material? Are you as a department, and then are you as an individual making sure that you're teaching as much as you can, as quickly as you can, and as deeply as you can? It's always that thing of coverage or depth, you know.* (Assistant Principal, Urban High School, CA)

But despite this recognition of positive potential, educators more often expressed concerns about unintended, ancillary impacts on students, teachers, and schools. They were especially concerned about the effects of accountability on student and staff motivation and their mission, as well as the technical details or fairness of existing or proposed accountability designs.

Maintaining high school students' motivation to learn and remain in school was the number one challenge mentioned by staff in a majority of our high schools. Teachers in a few schools thought state tests' rewards or sanctions could improve students' motivation to achieve. In Michigan, the scholarship dollars attached to the 11<sup>th</sup>-grade MEAP were said to have helped some students take the exam more seriously, but most thought this was a

sweetener for those already capable of pursuing postsecondary options but lacking the financial resources. One Pennsylvania teacher observed that test pressures had raised the importance and value of the high school diploma for many of her students. Test pressure had served as a bit of a “wake-up call” and had given her “more backing as a teacher” because now it was not just her saying students need to master the material but also “the state” (Teacher, Lakewood High School, PA).

But educators in states that attached strong sanctions for students to the tests, such as promotion and graduation, worried that they had harmed the morale of the lowest achieving students, and possibly spurred more to drop out. In one North Carolina school, for example, the assistant principal was deeply concerned about the “lie about dropouts”—the fact that the rate of 7–8% is reported to the press, but in actuality they regularly lose half of their freshmen class. He thought the EOC tests contributed to a downward spiral of failure. The assistant principal explained the scenario:

*These kids come in, and they have to pass this test to get through the course. They get a little behind and a little further behind, and they look, and they say, “I’m never going to pass this test,” and they drop out. (Assistant Principal, Lincoln High School, NC)*

Similarly, a math teacher at Grant High School in North Carolina, who generally thought that the EOC strategy helped students to maintain their grades also said:

*[But] I’m not a hundred percent supportive of all the students being held accountable, and say they’ve got to pass Algebra 1 to graduate from high school. Some of them can’t do it. Do we lose that kid to dropout just because he’s frustrated with the requirement, or do we teach him a trade that he can be productive in society?*

Michigan educators thought that testing geometry was too much to ask of all students, particularly the non-college-bound. Similarly, teachers from River City High School in New York thought that high-stakes testing aggravated student apathy because it forced a diet of irrelevant curricula:

*The new mandates coming down from the state make me crazy. On the one hand, we’re supposed to give kids authentic learning experiences, but more and more we’re forced to teach to a test. They don’t translate to anything real meaningful in their lives. (Teacher, River City High School, NY)*

Counter to policy intent, Florida educators felt that students did not care how they did on the exams required for graduation because they were permitted to retake the tests so often. Policy churn could also lower teacher and student motivation.

California educators anticipated that the state would postpone or cancel the use of CASHEE as an exit exam,<sup>3</sup> leading teachers in Urban High School to dismiss the exam. In addition, some California teachers perceived that students did not care much about the state Standardized Testing and Reporting (STAR) test either:

*But it’s definitely true that I’ve had classes where we’ve passed out the STAR test and we’ve had kids bubbling C all the way down and put their head down for the rest of the class, and multiple times. (Teacher, Arnold High School, CA)*

---

<sup>3</sup> Indeed, after we were in the field, the State Board of Education postponed the effective date of using the CASHEE as an exit exam from 2004 to 2006, to ensure that students had an adequate opportunity to learn to the standards.

Teachers suspected that the large number of California state tests lowered students' motivation to do a good job on them.

The State Board of Education in North Carolina had many discussions about appropriately balancing pressures on schools and teachers versus students. Their decision to make the EOC tests one quarter of the final grade was an attempt to respect teacher judgment but also to give students an incentive to take the EOC seriously.<sup>4</sup> Nevertheless, educators in one struggling North Carolina school (Lexington High School) said that students did not care enough about the EOC tests<sup>5</sup> because they could still pass their courses without doing well on the test. But since the public relies on the EOC results to make judgments about the school, Lexington teachers felt considerable stress. The principal suggested that good teachers had quit as a result of the pressure, and staff morale was perceived to be low in other North Carolina schools. The morale problem is compounded in low-performing schools that lack a cohesive professional community. For instance, in Urban High School in California, where teachers did not regularly communicate about instruction and dialogue was infrequent, the accountability system seemed to reinforce low staff morale:

*Well, there's always, it always puts pressure on you. . . . The public's putting pressure on you, and saying you're not good enough. . . . In a lot of cases, it has a deleterious effect on how people teach . . . it's demoralizing. You're doing everything you can, you're working as hard as you can with the students you have sitting in front of you, recognizing where they are and trying to pull them along as fast as you can to where*

---

<sup>4</sup> Correspondence from Charles Thompson, Professor of Education at East Carolina University (personal communication, October 18, 2004).

<sup>5</sup> Note that we did not interview teachers responsible for the eighth-grade basic skills test, which is required for graduation.

*you'd like them to be. The implication that nobody had any standards before they came up with these things is insulting. (Teacher, Urban High School, CA)*

As with testing, many perceived that the design of state accountability had narrowed their traditional missions. While state accountability programs include student dropout rates and other measures, testing is given the largest weight in determining schools' progress. Educators in our schools thought that the measurement of their success in these terms had damaged their traditional vocational missions, special emphases on the performing arts, or other goals. For instance, the principal in Roberts High School in Michigan implied that the devaluation of their vocational mission had harmed their strong suit as a school, as well as students' employability after graduation.

Finally, educators from many schools viewed their accountability measures, particularly the heavy reliance on subject matter tests, as an unfair and illegitimate way to judge the success of their school. Sample schools reported that many of their freshmen entered with very low reading abilities and with many other academic and social problems. Said one Florida math teacher:

*It's really not fair to compare the school with other schools, if you know what I mean. If they switched—like, what they call A schools, if they switched out the faculty here with the Fs, with the D-school faculty, I know they wouldn't bring it up to an A and we wouldn't bring them down to a D. You just have different clientele, you understand? (Math Teacher, Oceanside High School, FL)*

They felt that students were so far behind in skills that the state tests were not a good measure of the progress they had made with

those students (e.g., Nelson High School, NY). For some schools, simply getting students to attend class was an important success, one not accounted for by their state’s emphasis on testing. Finally, administrators and teachers also expressed concerns over the fairness and accuracy of comparing their results to those of other schools, particularly when funding disparities between schools remained, or since schools served students with very different socioeconomic backgrounds.

## **Salience of State Standards, Testing, and Accountability**

We anticipated that educators’ perceptions about the value of the state system would give impetus to—or, if negative, depress—the pressure they felt from and/or the extent to which they responded to their state accountability system. While this held true in some specific instances, we found major disconnects between educators’ views on the merits of the state system and its salience in high schools.

For example, while state content standards had long been a fixture of the policy environment in California and were perceived as highly legitimate, one of the California schools in our sample was only just beginning to align its curriculum to standards. A number of California schools did not have any processes in place to secure alignment to standards, and implementation was reportedly mixed. We found a similar situation in some of our Florida schools. At Atlantic High School, the principal—an advocate of standards—thought veteran teachers were resisting the integration of standards into the school’s curriculum, and had to wait for retirement to build a staff more willing to buy into standards. As this suggests, leadership was crucial. For instance, district leadership greatly enhanced

the prominence of standards in Lakewood and Mountain View High Schools in Pennsylvania.

In North Carolina, the design of the accountability system gave standards substantial clout. There, course-based standards and course-based tests provided educators with strong guidance and motivation to use that guidance. This “seamless system” also fit nicely into the organizational structure of high schools, because it did not require as much work for high school teachers to determine who would teach what knowledge and skills as in states that had benchmarked standards without more specific guidance documents.

In contrast to standards, and despite a high volume of criticism and concern, state tests were most likely to generate pressure and high levels of activity. Test-related initiatives ranged from more ancillary test preparation activities to the adoption of whole courses specifically designed to address skills and content on state tests. The response in New York best illustrates the point that negative views about testing did not necessarily depress school level action. New York educators expressed angry challenges to the design and impact of their state testing and accountability programs; indeed the amount of criticism there was notably higher than in other states. Nevertheless, schools undertook a significant amount of action to address measured performance, above and beyond such state mandates as remediation. Educators in Nelson High School, for example, thought the Regents exams were not good measures of students’ content knowledge. But even though the school was listed as “satisfactory” on the state accountability index, test-related action was abundant: the English and mathematics departments consistently included Regents questions in classes, conducted test preparation and remedial activities, used state test data, and aligned their textbooks as well as summer reading lists to state standards and tests. Teachers also reported raising student course grades if students passed the Regents exams.

In Michigan, schools with the highest volume of criticism were the most responsive in their state context. For example, Roberts High School staff argued that MEAP and standards were not well aligned, did not map well onto students' abilities and development, and had sharply devalued their vocational mission. Nevertheless, they systematically planned changes in instruction to respond to MEAP, including research paper assignments for MEAP, test-taking skills seminars, practice tests, MEAP skills in student planners, and curriculum alignment. Interestingly, compared to several other Michigan schools, the staff we interviewed met and talked frequently to one another about professional matters. Teachers also expressed a high level of trust for their school administrators. This apparently strong professional culture may explain the unusually high level of response here and in another Michigan school.

Many administrators and teachers talked about using state and other test data to make changes in their curriculum, consider student placement, and the like, although use appeared to vary widely from school to school, and state to state. Some teachers and even principals in the weakly responsive Michigan schools admitted that they had never seen MEAP results, or had only read about them in the newspapers. By contrast, the majority of the teachers we interviewed in North Carolina looked at and used the results of their students' EOC exams to target areas where they as individuals needed to improve. For example, one math teacher in Grant High School noted:

*One of the first strategies that I use is from year to year, I take that testing data, which ...breaks down those objectives into specific objectives. And I can look at my class's achievement for those individual objectives, and I can pinpoint areas that I need to improve. I take my lowest objectives and those are the ones I make modifications in the structure. One year it was radicals. And*

*so the next semester I did a lot more with explanation of radicals.*

The mathematics department in that same school used the results of the EOCs to realign their curricula, particularly in Algebra 1A and 1B, and to require that students earn a grade of C before taking the next level of mathematics. The gulf in data use between a state like Michigan and North Carolina is explained by the fact that in Michigan, state test results are released once a year, but only for (primarily) 11<sup>th</sup>-grade students, and until recently the state provided no item skills analysis. Some Michigan educators thought the MEAP data were not trustworthy or useful for diagnostic purposes. In North Carolina, the data were course-specific and timely, and were often used by individual teachers.

While high-stakes environments did generate more press in general, we found that even there the vast majority of high school teachers did not experience any direct consequences for success or failure, and most did not think that administrators really knew what was going on in their classrooms. Administrators or department chairs in many schools collected lesson plans and required teachers to document curriculum alignment to standards; in some cases, districts or states (e.g., Michigan) mandated these activities. However, more often than not, teachers did not receive feedback on these lesson plans, and teachers' instruction was rarely monitored or evaluated outside formal tenure and evaluation requirements. Department chairs did not have the power or authority to observe teachers, remove teachers, or mandate instructional change, and those who did foray into changing their colleagues' instructional practice did so delicately, trying to mask their efforts. As a result, even in high-stakes environments like Florida, teachers did not feel as though anyone was holding them accountable for the performance of their students or their school. Teachers' work remained largely uncoupled

from the system of sanctions and rewards. Said one California teacher:

*No one's ever come in here and said, your test scores are too low, what are you doing about it? If you don't improve, you're going to lose your position here . . . we're working in a really difficult neighborhood with a difficult student population, but when, hmm, but when it comes down to it, honestly, there . . . I don't feel threatened that I'm going to lose my job if enough of my students don't pass the high school exit exam when they get to be a senior.* (Teacher, Arnold High School)

Nevertheless, teachers in both weak and strong accountability states articulated feelings of responsibility for their students' test scores, as well as other academic outcomes—getting their students ready for the next course level in their subjects, and postsecondary futures (i.e., college or work). Teachers' feelings of responsibility about test results tended to come from their own sense of professional obligation, concern for student success, responsibility to their colleagues or community, and/or the articulated concerns of their principals about school test results. For the most part, principals reported much higher levels of stress about testing results than their teachers, and communicated these concerns to their staff but did not actively manage incentives or day-to-day instruction to meet these goals.

In addition to the pressures of student achievement outcomes, principals in some of our high schools questioned staff about course failures. Attendance and student dropouts were acknowledged concerns in these buildings; in one Michigan school (Jones High School), for example, the principal noted freshman failure rates of 50%. Some teachers said their principals took them to task on these numbers, and felt that they had to justify these grades, or alternatively to find ways to help these students pass. One teacher complained

that students in her “so-called” required senior course were allowed to substitute other credits to graduate. This illustrates the dual and sometimes competing goals of accountability, which measures school success both on achievement and attainment measures. In educational practice, these goals often conflict and pose a tough challenge to high school teachers and administrators.

We expected that we would find more consistent levels of response in high-stakes accountability environments, but this was not borne out. Instead, we found considerable variation in the salience of the system even in those strong accountability states, regardless of schools' relative location on the accountability index. So, for instance, while our lowest performing schools were more responsive to external accountability than some research would have predicted given their circumstances (see Elmore, 2003; O'Day, 2004), we also encountered some very low-performing schools where addressing state accountability was a distant concern. For example, teachers in two of the three priority schools in our North Carolina sample reported feeling less press to address accountability than teachers in the third school.

The evidence about the salience of the accountability system leads us to five main points:

1. Examples of the disjuncture between acceptance of the components of accountability and levels of action in high schools suggests that *the press of the system was getting through, despite significant doubts and concerns about the impact on teaching and learning*. Response without conviction runs contrary to the conclusions of a body of implementation literature which states that action occurs after “street-level bureaucrats” become

- committed to and engaged with the reform agenda (Lipsky, 1980; McLaughlin, 1990).
2. While the consequences of accountability failures or successes on two actors in the schools—principals and students—were usually clear and direct, *the consequences for teachers were not well articulated*. (See also Goertz, 2001.)
  3. Nevertheless, teachers did focus on academic outcomes out of a sense of professional responsibilities that may have been accompanied by perceptions of more informal pressures from their administrators, colleagues, or the community.
  4. We found substantial differences in the extent to which state testing was salient to schools in the strong accountability systems of CA, FL, NC, and NY compared to schools in the weaker systems of MI and PA. While schools in the latter two states did address tested performance, their accountability systems were simply not as salient or pressing. This occurred even though the public reporting of MEAP scores had been part of the policy scene for many years—pointing out that often, public reporting is not enough to focus attention (for contrasting findings on the effects of public reporting, see Goldhaber & Hannaway, 2004). Certainly states with higher stakes had a legal obligation to students to provide more test-related opportunities and services. But again, these legal obligations did not explain all the types of responses occurring in these high schools.
  5. Despite this general pattern of higher salience in the high-stakes environments, it is extremely important to point out that the level of response within states was not consistent. In fact, *we found as much variation in strong accountability states as weak ones*. It is to this issue that we turn in the next section of the chapter.

## Explanations for Variation in Schools' Response to External Accountability: The Pull From Below

Several factors seem to account for the wide variation in high schools' response to external accountability. District-level leadership was critical, as were other local contextual factors such as community press, perceptions that ultimate consequences were a realistic possibility for the school, the school's professional culture, teachers' feelings of efficacy, and the school's capacity to respond to accountability challenges.

Active district leadership supporting accountability was associated with greater high school response inside all the strong accountability states, and even in Pennsylvania, with its weak and less stable system.<sup>6</sup> For instance, when the Renaissance City School District was placed on the state warning list for poor performance, the long-time superintendent decided that his earlier approach of delegating

---

<sup>6</sup> With rare exceptions (Hampton City and to a modest degree Foggy Mountain City), the districts in our Michigan sample were not very active vis-à-vis intervening in their high schools. During our fieldwork, public reporting of MEAP results was the only consequence in place for high schools or districts.

school improvement initiatives to the schools had failed. He started to recentralize control over the schools, and held them accountable for raising PSSA scores. He said, “The new number one job of administrators is improving student performance in reading and math.” The Lakewood High School principal in turn made the PSSA his primary focus and held his staff accountable for test scores. He met with teachers to discuss PSSA performance both at the whole school and department levels. In addition, he required teachers to use targeted PSSA workbooks. An English teacher said:

*I can also feel the pressure that is on [the principal]. Because when pressure is on him from the superintendent, which is an incredible pressure, I can hear it through what he says. You know, not directly, but I can hear that pressure. And so it really keeps me alert and on my toes that, you know, you really need to be doing what you can to help him, which ultimately helps the scores. (Teacher, Lakewood High School, PA)*

Even low-capacity districts could stimulate higher levels of focus. For instance, one small North Carolina district with few central office staff used test score data to monitor school progress and signal the importance of student performance. The district also directed resources where needed, particularly to the lowest performing schools.

District leadership could stimulate school action even when schools were relatively high-performing. For example, because its district was highly focused on test results and accountability, staff in relatively well-performing Southern High School (NC) expressed great fear about slipping into a lower rating. They discussed a variety of ways they used the EOC exams to drive their practice, including using test data to identify conceptual gaps and target students for extra help.

Geography and policy design, along with district size, could influence district intervention on behalf of accountability. In Florida, for instance, districts focused on low-performing schools more than those in higher accountability categories, in part because state law required them to provide special assistance to these schools. But Florida district staff were stretched thin by the large number of schools in their jurisdiction. County administrators in two of our sample districts had 30 and 38 high schools, respectively. Even the smaller rural Florida counties managed five to six high schools—the same number of schools as our largest district in Michigan. The confluence of these factors led educators in our middle-performing Florida schools to respond less actively to state accountability.

But district press was not in and of itself a sufficient factor in schools’ engagement with external accountability; we visited several schools that resisted their districts’ leadership in this area. One of our California districts targeted three schools performing poorly on state measures. In addition to a stronger press for improvement, the district provided more professional development and support to these high schools. But this pressure and support did not seem to erase the view in one of the target schools, Arnold High School, that no consequences were really likely to befall them or their students. They did not feel much pressure from their poor performance. Staff at Urban High School also expressed this sentiment:

*The adults say a lot of things that never really happen, like you’re not going to pass, you’re not going to graduate unless you take this test . . . but somehow these kids wiggle and worm and it all sort of falls into place . . . there’s going to be whatever it is going to come along and sabotage that exit exam. (Teacher, Urban High School, CA)*

For Arnold High School, the lack of responsiveness to their district's press on external accountability was an issue of staff's knowledge and feelings of efficacy. The principal reported that while he did not know how to meet the goal that all students could succeed at high levels, his strategy was to emphasize this belief and ask his staff what they needed to make it happen. Yet he was turning to a staff of young and inexperienced teachers: Only 66% were certified, and many were teaching on emergency credentials. The school's Academic Performance Index (API) statewide score had not been above a 2 (toward the bottom of the accountability spectrum) since 1999, and staff had come to believe that they would never be able to change these dismal results.

On the other hand, perceptions of efficacy at the school level could also produce resistance to district leadership. For example, Medal County, North Carolina, focused heavily on data and the EOC exams. Although their initiatives influenced practice in Lincoln High School (teachers reported the exams' strong impact on their instruction, goals, efforts to search for improvement strategies, and course assignments), it did not have much effect in neighboring Maple High School. In this school, only the math department reported using the EOC scores to place students in classes. Such differences appeared to be due to the fact that Maple performed well on the state system, and had a relatively new, forceful principal who sought to buffer the school from the external pressures exerted by the district. He felt the school was doing well and did not need to make substantial change. Further, staff thought that district efforts to measure and spur student achievement—like the state's—were *not* good indicators of student learning, and in fact had led to lower standards for both teaching and learning.

Community expectations about academic performance—or more precisely schools' perceptions of their community and its

expectations—appeared to be another important ingredient influencing the responsiveness to external accountability of some schools, like Redwood High School in California. Although this school had a high position on the API—especially compared to schools serving similar populations—its staff pressed extremely hard to improve their results. The school was located in a wealthy community with high expectations for its schools. In addition, it had committed school leadership, ample fiscal resources, and a strong collegial atmosphere, all of which contributed to a strong sense of internal accountability (see Debray, Parson, & Avila, 2003). Similarly, in Striver and Mountain View high schools in Pennsylvania, staff concerns about community perceptions' clearly drove their desire to improve performance on the PSSA.

## Conclusions and Research Implications

High school teachers and administrators were aware—often keenly aware—of the challenges posed by external accountability, and spoke of the unintended consequences that such systems could create for their teaching and curriculum or for student motivation and persistence in school. Many were deeply mistrustful of state tests or other technical aspects of the accountability design, and had serious doubts about whether testing and accountability, in operation, were working to improve their educational practice. Nevertheless, even the most skeptical acted to address the demands of their testing and accountability programs—indeed, sometimes these schools were the most responsive. On the other hand, though standards were well accepted in most states, they did not always lead to much focused action. While in part this has to do with the fact that some schools and districts had already addressed standards

reforms in prior years, others had never done so.

These findings challenge the longstanding notion that implementers' belief is a necessary prerequisite to action. Part of the explanation may lie in the fact that while educators questioned the immediate effects of testing and accountability, they also recognized its potential value to ultimately improve various aspects of schooling. In future research, simplistic notions that "belief follows action" should be replaced with more discriminating models of the relationship between these two domains.

Certain features of accountability designs raised the likelihood of action: Stronger consequences for students and schools did, in general, yield greater press and response, particularly when the system was stable. But there was a notable exception to this rule. Even though accountability consequences rarely had direct effects on teachers' employment, their professional identity, care and concern for students, feelings of efficacy about their ability to address the challenges posed by accountability, and concerns for their administrators and community coupled them to the goals of external accountability.

Furthermore, while we found confirming evidence that strong and stable accountability stimulated higher levels of press and action to improve on accountability measures, we also found that schools were not consistently responsive in any state system, be it weak or strong. Nor were schools consistently responsive if they were in a particular performance level in the system. We did learn that high schools tended to be more active when their district leaders were focused on accountability. The early RAND studies in the 1970s also identified districts as crucial players in school-level implementation, but this lesson was oft-forgotten in the policy world of the 1980s and 1990s. At that time, state and district education agencies were seen as anathema to improvement; popular reforms, such as site-

based decision making, were designed to bypass districts and give schools greater autonomy from their bureaucracies. The wisdom of harnessing districts to the reform agenda has been lately rediscovered, but it is a lesson worth reiterating in the name of improving the efficacy of accountability. Understanding why some districts press for high schools to address accountability while others do not—and how they can do so more effectively in high schools—is an important question for researchers and policymakers to continue to pursue.

Finally, educators were strongly concerned about the narrowing of the traditional mission of high schools, as well as various kinds of curricular and instructional narrowing that they perceived. To be sure, a central goal of the standards reform movement has been, in fact, to rein in the extremely diverse high school curriculum and to pare down nonacademic courses. The comprehensive high school curriculum began to emerge in the early part of the twentieth century when Progressive educators decided to go beyond the classics to prepare the greater majority of students for their certain futures: marriage and motherhood for girls and work in the new factories for boys. Curricular differentiation in high school expanded during the 1960s, when educators tried to make the curricula more socially relevant and engaging, leading to the much-criticized "shopping mall" curriculum (Powell, Farrar, & Cohen, 1985).

Efforts to bring a more academic focus to the high school curriculum began in the early 1980s, with the landmark *A Nation at Risk* report, and extended to standards-based reforms accompanied by performance accountability. It is clear, however, that these efforts push against some deeply held values about the role of high school. As Leslie Siskin once wrote:

*High schools . . . are being asked to take on a new task—something they were not*

*designed to do—to prepare students for a defined minimum academic standard, and to get all students to graduate by achieving that standard. We have certainly not organized high schools so that all students would take the same content, or meet the same standards to graduate. In fact, comprehensive high schools were historically designed to do precisely the opposite . . . their design imperative has been to serve democratic purposes and accommodate diverse student populations by creating a wide range of programs, and a differentiated curriculum. (Siskin, 2003, pp. 176–177).*

Our educators' concerns about this more restrained academic focus reflect a deep disagreement or at least consternation about whether it can engage and better educate all students. Their comments also indicate that narrowing occurs in a variety of ways—some in the spirit of reform ideals, some not. The Florida school that eliminated special projects to allow more time for academic courses seems to meet reform goals. But when students retake the same courses over and over until they pass state tests, the academic purpose seems to be lost. Researchers should make a closer study of just what kind of narrowing is occurring, and how it may be influencing students' persistence in school.

**Table 1. Key Characteristics of State Accountability Policies for High Schools, 2002–2003:  
Six Study States**

|    | Target of accountability | Assessment                             | Accountability Measure                                                                                                                                             | Consequences                                                                                                                   | Stability       | Strength of overall system <sup>a</sup> |
|----|--------------------------|----------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|-----------------|-----------------------------------------|
| CA | Student                  | CAHSEE: grade 10                       | Students must score 60% in ELA and 55% in math.                                                                                                                    | Graduation: class of 2006; scholarship money.                                                                                  | Unstable        | Strong (4)                              |
|    | School                   | CA Standards Test (CST), CAHSEE, CAT/6 | Academic Performance Index (API): based on performance and growth.                                                                                                 | Monetary rewards for growth. Sanctions for low performers: TA, outside intervention, and possible takeover.                    | Stable/Emerging |                                         |
| FL | Student                  | FCAT: grade 10                         | Passing score is middle of Basic (Level 2 of 5).                                                                                                                   | Graduation: class of 2003; Certificate of Achievement for higher score.                                                        | Stable/Mature   | Strong (5)                              |
|    | School                   | FCAT: grade 10                         | A+ Plan: based on performance, growth overall, and gains of lowest performing students.                                                                            | Monetary rewards for high performance and/or growth. Sanctions for low performers: TA, student choice, and reconstitution.     |                 |                                         |
| MI | Student                  | MEAP: grade 11                         | Performance level is Basic.                                                                                                                                        | Diploma endorsement.                                                                                                           | Stable          | Weak (1)                                |
|    |                          |                                        | Performance level is Proficient.                                                                                                                                   | Scholarship money.                                                                                                             |                 |                                         |
|    | School                   | MEAP: grade 11                         | None.                                                                                                                                                              | None (rewards?).                                                                                                               | Unstable        |                                         |
| NY | Student                  | Regents Comprehensive Exams (RCEs)     | Passing score of 65 (Proficiency) or higher for students entering ninth grade in 2001–2002; local option to set score at 55 (Basic Proficiency) for prior classes. | Graduation: class of 2000.                                                                                                     | Stable/Emerging | Strong (5)                              |
|    | School                   | RCEs                                   | Performance Index (sum of percentage of students scoring above Basic Proficiency and percentage above Proficient).                                                 | Rewards for high performance and/or growth (?). Sanctions for low-performers: TA, additional funds, and loss of accreditation. |                 |                                         |
|    | District                 | RCEs                                   |                                                                                                                                                                    | May be designated Below Standards; develop improvement plan.                                                                   |                 |                                         |

|    | Target of accountability | Assessment                                      | Accountability Measure                                                                                                                | Consequences                                                                                                                                          | Stability       | Strength of overall system <sup>a</sup> |
|----|--------------------------|-------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|-----------------------------------------|
| NC | Student                  | 8 <sup>th</sup> Grade End-of-Grade (EOG) Exam   | Passing score: achieving on grade level.                                                                                              | Graduation.                                                                                                                                           | Stable/Mature   | Strong (5)                              |
|    |                          | End-of-Course (EOC) Exams                       | Passing score.                                                                                                                        | 25% of course grade.                                                                                                                                  |                 |                                         |
|    | School                   | EOC; 10 <sup>th</sup> -grade Comprehensive Test | Percentage of students passing EOCs; expected growth on EOCs, between 8 <sup>th</sup> - and 10 <sup>th</sup> -grade competency tests. | Monetary rewards for higher performance and/or growth. Sanction: state assistance team, removal of principal.                                         |                 |                                         |
|    | District                 |                                                 | Half of schools are in low-performing category.                                                                                       | SDE can replace superintendent or other administrators; LEA can lose accreditation.                                                                   |                 |                                         |
| PA | Student                  | PSSA: grade 11 or Local Assessment              | Proficiency on PSSA or local assessment as determined by LEA.                                                                         | Graduation.                                                                                                                                           | Unstable        | Weak (1)                                |
|    | School                   | PSSA: grade 11                                  | Increase PSSA scores by at least 50 points, increase attendance.                                                                      | Monetary rewards.                                                                                                                                     | Stable/Emerging |                                         |
|    | District                 | PSSA                                            | 50% or more of students at Below Basic performance.                                                                                   | Monetary rewards for improving performance of ELL, disabled, poor students. Sanctions for low performance: DIP, TA, additional funds, state takeover. |                 |                                         |

*Note:* A Rating of “Weak” to “Strong” determined by authors on the basis of target of accountability and strength of consequences. Numeric rating ( ) assigned by Carnoy and Loeb (2004). States assigned a rating of 1 have state assessments but no school or student sanctions. States assigned a rating of 5 test students in multiple grades, strongly sanction and reward schools, and require students to pass a high school graduation test. States with strong school sanctions but no high school exit exam are assigned a rating of 4.

**Table 2. Characteristics of State Assessment Systems for High Schools, 2002–2003: Six Study States**

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CA | <p><i>CAHSEE (Grade 10)</i> in ELA and mathematics. ELA aligned to 9<sup>th</sup>/10<sup>th</sup>-grade standards; mathematics aligned to sixth- to eighth-grade standards, including Algebra 1. Multiple choice format (two writing items). Students are tested in March or May. Students who fail have up to three retakes. New test implemented in 2000–2001.</p> <p>State funds summer school for students in grades 7–12 in danger of failing the exam. LEAs must provide remediation and supplemental instruction to students who fail.</p>                            |
|    | <p><i>California Standards Tests (CSTs)</i> in English, mathematics, and science in grades 9–11 and in history/social science in grades 10–11. Aligned to state standards.</p>                                                                                                                                                                                                                                                                                                                                                                                               |
|    | <p><i>CAT/6</i> in reading/language skills, math and science in grades 9–11. Norm-referenced test (NRT). Becoming smaller component of state assessment system.</p>                                                                                                                                                                                                                                                                                                                                                                                                          |
| FL | <p><i>FCAT (Grade 10)</i> in ELA and mathematics. Standards-based and norm-referenced items. Standards-based mathematics items cover algebra and geometry and are aligned to Sunshine State Standards. NRT mathematics topics include algebra, geometry, trigonometry, and precalculus. Mixed item format. Test is given in March. Students may retake exam five times in grades 11 and 12. Replaced High School Competency Test in 1998.</p> <p>LEAs must provide remediation for students who fail exam. State provides Supplemental Academic Instruction (SAI) funds.</p> |
| MI | <p><i>MEAP (Grade 11)</i> in ELA, mathematics, social studies, and science. Criterion-referenced tests. ELA has writing, reading and listening component. Mathematics covers algebra and some geometry. Test is given in January. Students may take in 10<sup>th</sup> grade and retake in 12<sup>th</sup> grade. Revised in 2001–2002 (math) and ELA (2002–2003) to align with 1996 Curriculum Frameworks. Replaced High School Competency Test.</p> <p>Remediation is district/school option.</p>                                                                          |
| NY | <p><i>Regents Comprehensive Exams (RCE)</i> in English, Mathematics A for class of 2001. Subsequent classes must take two additional tests in social studies and one in either science or a foreign language. Students take at completion of course. Are aligned to state Learning Standards and to grade 9–12 courses of study. Students may take component retests for sections they failed.</p> <p>Schools must provide Academic Intervention Services (AIS) to students at risk of not passing exams; LEAs design programs.</p>                                          |
| NC | <p><i>8<sup>th</sup> Grade End-of-Grade (EOG) Exams</i> in reading comprehension and mathematics. Aligned to state’s eighth-grade Standard Course of Study. Replaced Minimum Competency Test as graduation requirement in 1994. Students may retake exam through 12<sup>th</sup> grade.</p> <p>Schools must provide remediation to students who fail eighth-grade EOG exams. Remediation for EOCs is local option and participation is voluntary on part of student.</p>                                                                                                     |

|    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| NC | <p><i>End-of-Course (EOC) Exams</i> in 11 high school courses ranging across 9<sup>th</sup> to 12<sup>th</sup> grades, including Algebra I; Algebra II; Biology; Chemistry; Economic, Legal, and Political Systems; English I; English II; Geometry; U.S. History; Physical Science; and Physics. Aligned to state’s course specific Standard Course of Study. Students take at completion of course. In place since mid-1980s. Became part of state accountability system in 1997–1998.</p> |
|    | <p><i>10<sup>th</sup> Grade High School Comprehensive Test</i> in reading comprehension and mathematics designed to measure growth from 8<sup>th</sup> to 10<sup>th</sup> grade. Multiple choice format. No student consequences.</p>                                                                                                                                                                                                                                                        |
| PA | <p><i>PSSA (Grade 11)</i> in mathematics, reading, and writing. Writing also tested in ninth grade. Aligned to state standards. Mathematics coverage can go through calculus. Mixed item format. Students are tested in late March and can retake PSSA in 12<sup>th</sup> grade.<br/>Remediation is local option.</p>                                                                                                                                                                        |

## References

- Abelmann, C., & Elmore, R. F. (with Early, J., Kenyon, S., & Marshall, J.). (1999). *When accountability knocks, will anyone answer?* (CPRE Research Report No. RR 42). Philadelphia: Consortium for Policy Research in Education.
- Berman, P., & McLaughlin, M. W. (1975). *Federal programs supporting educational change. Volume 4: The findings in review.* Santa Monica, CA: RAND.
- Carnoy, M., & Loeb, S. (2004). Does external accountability affect student outcomes? In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 189–219). New York: Teachers College Press.
- Center on Education Policy. (2003). *State high school exit exams: Put to the test.* Washington, DC: Author.
- Debray, E., Parson, G., & Avila, S. (2003). Internal alignment and external pressure. In M. Carnoy, R. Elmore, & L. S. Siskin (Eds.) *The new accountability: High schools and high-stakes testing* (pp. 55 – 85). New York: RoutledgeFalmer.
- Elmore, R. (2003). Accountability and capacity. In M. Carnoy, R. Elmore, & L. S. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 195–209). New York: RoutledgeFalmer.
- Elmore, R. F., Abelmann, C., & Fuhrman, S. (1996). The new accountability in state education reform: From process to results. In H. F. Ladd (Ed.), *Holding schools accountable* (pp. 65–98). Washington, DC: Brookings Institution.
- Fuhrman, S. H. (1999). *The new accountability* (CPRE Policy Brief No. RB-27). Philadelphia: Consortium for Policy Research in Education.
- Goertz, M. E., & Duffy, M. C. (2001). *Assessment and accountability systems in the 50 states: 1999–2000* (CPRE Research Report No. RR-046). Philadelphia: Consortium for Policy Research in Education.
- Goertz, M. E. (2001). Standards-based accountability: horse trade or horse whip? In S.H. Fuhrman (Ed). *From the capitol to the classroom: Standards-based reform in the states. The one hundredth yearbook of the National Society for the Study of Education: Part two.* Chicago: National Society for the Study of Education.
- Goldhaber, D., & Hannaway, J. (2004). Accountability with a kicker: Observations on the Florida A+ Accountability Plan. *Phi Delta Kappan* 85(8), 598–605.
- Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public services.* New York: Russell Sage Foundation.
- Massell, D. (1994). Achieving consensus: Setting the agenda for state curriculum reform. In R. F. Elmore & S. H. Fuhrman (Eds.), *The governance of curriculum* (pp. 84–108). Arlington, VA: Association for Supervision and Curriculum Development.
- Massell, D. (2000). Developing innovative content standards in the fields of mathematics and social studies (Doctoral dissertation, Stanford University, 2000). Dissertation Abstracts International, A61/09, 3455.

McLaughlin, M. W. (1990). The Rand Change Agent study revisited: Macro perspectives and micro realities, *Educational Researcher*, 19(9), 11–16.

O'Day, J. (2004). Complexity, accountability, and school improvement. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 15–43). New York: Teachers College Press.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The shopping mall high school: Winners and losers in the educational marketplace*. Boston: Houghton Mifflin.

Siskin, L. S. (2003). When an irresistible force meets an immovable object. In M. Carnoy, R. Elmore, & L. S. Siskin (Eds.) *The new accountability: High schools and high-stakes testing* (pp. 175–194). New York: RoutledgeFalmer.

Siskin, L. S. (2004). The challenge of the high schools. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 167–188). New York: Teachers College Press.

U.S. Department of Education. (1994). *What do student grades mean? Differences across schools*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.