

Executive Summary

In June of 1998, the California Department of Education (CDE) awarded a contract to WestEd, in collaboration with Management Analysis and Planning, Inc. (MAP), to evaluate standards-based accountability in California. The evaluation was designed to examine the state's Standards-Based Accountability System and its relationship with other accountability efforts and initiatives. In particular, the evaluation focused on the status and impact of local accountability systems, content standards, assessment measures, use of data, consequences and incentives, and challenges and assistance in school districts across the state.

California's Standards-Based Accountability System was established in 1997 by CDE in response to federal requirements of Title I of the Improving America's School Act (IASA), and in accordance with the state's IASA plan. The Standards-Based Accountability program was largely a system of locally defined programs. Districts were expected to develop and adopt local content and performance standards (comparable to state standards) and to use multiple assessment measures to determine whether students were meeting the local standards. As part of the state's Consolidated Application for Funding Categorical Aid Programs, Part II, districts were required to report on the progress of their standards-based accountability systems. In particular, districts were to include a description of their accountability and assessment system along with student achievement data for all schools within the district.

Significant events occurred during the study which have some bearing on its implications. In April of 1999, the legislature enacted and the governor signed a new state-driven accountability program, the Public Schools Accountability Act (PSAA) of 1999. This law created three basic components, which together constitute a new statewide accountability system: 1) the Academic Performance Index (API), an index to rank the performance of schools; 2) the Immediate Intervention/Underperforming Schools Program (II/USP), an assistance and intervention program for schools who fall below expectations; and 3) the High Achieving/Improving Schools Program (HA/ISP), a rewards program for schools who exceed them.

The system created by the PSAA essentially replaced the program that this study was in the midst of evaluating. Districts still have the option to continue developing their own standards-based systems, but they will no longer be required by the state to use multiple measures or to submit accountability reporting information as part of the Consolidated Application. This change in law meant that, as data were being collected from the field on accountability, much of the information reported by district and school personnel reflected not only the progress to date on the pre-existing program, but also reactions to the PSAA.

Research Questions

Ten major research questions served as the focus of the evaluation of the Standards-Based Accountability system:

1. At what stage or level are districts in planning and implementing their standards-based accountability systems?
2. What is the nature of local content and performance standards in language arts and mathematics and how do they compare with the state standards? What mechanism does the district use to compare its standards with the state's model standards?
3. To what degree have district standards been implemented in schools? At what stage or level are districts and schools in planning and implementing the district standards?
4. What is the nature of local assessments for standards-based accountability and how are they used to determine whether a student has met or has not met the local standards?
5. How are data from local accountability systems used?
6. What types of performance targets do districts set for schools?
7. What types of incentives do districts provide for schools to meet their targets? What consequences do schools face if they do not meet the targets?
8. What practices or features of a district's standards and accountability system are associated with particular educational outcomes and practices?
9. What practices of the state education agency and other education-related institutions and assistance centers help or hinder districts in implementing their standards and accountability systems?
10. What obstacles do districts face in implementing a standards-based accountability system and how can the state education agency and other educational institutions and assistance centers help districts in overcoming those obstacles?

Conceptual Framework

In addition, the research team developed a conceptual framework as a lens through which to view accountability. The conceptual framework designed by the research team includes

broad elements of an “ideal” accountability system. In particular, the conceptual framework has six interrelated and reinforcing elements:

- Alignment of state and local content standards
- Student performance standards and aligned assessments
- Ongoing data analyses and reviews of school performance
- School improvement and intervention strategy
- Stakeholder involvement and engagement
- Continuous improvement of an accountability system

Methodology

The school district constituted the primary unit of analysis for the evaluation. Researchers also focused some attention on the state and school levels, primarily through interviews. Evaluators employed a mix of quantitative and qualitative methods and a variety of data collection instruments. The evaluation questions and the conceptual framework served as the basis for design of the data collection instruments.

An extensive survey of 200 California schools districts about their accountability practices formed the cornerstone of the evaluation. The sample of 200 districts included the state’s 120 largest districts, 50 “medium-size” districts, and 30 “small” districts; 133 districts (66.5%) returned their questionnaire. The questionnaire, administered in the spring of 1999, included items on the following topic areas: local accountability system features, content standards, student assessment measures, analysis and use of data, review of school performance, and helps and hindrances.

Interviews conducted at several levels formed another key component of the study. At the state level, interviews were conducted with several top-level policymakers, including legislators, legislative staff, and officials from the Governor’s Office and the State Board of Education. In addition, information was obtained from five high-level “policy implementers” at the California Department of Education. These interviews took place early in 1999.

At the district level, members of the research team interviewed personnel in each of seven school districts during the spring of 1999. The interview protocol included questions about accountability in general, standards, timelines, the roles of various actors in the accountability system, the use of data, challenges faced, successful practices, sources of assistance, and the new state accountability system. Research staff also spoke with representatives from about 30 small districts at a conference in March 1999.

Project staff were able to leverage school-level visits from another CDE-sponsored study, the Mathematics Implementation Study, also being conducted by WestEd/MAP. In the spring of 1999, trained mathematics observers interviewed and visited the classrooms of 55 teachers in eight California school districts; 28 were fourth-grade teachers and 27 were eighth-grade teachers. The principals at most of the schools also were interviewed. Although the interviews and classroom observations centered on matters related to mathematics instruction, they also touched on accountability, standards, and assessment.

Other data collection instruments used in the evaluation included the following: a review of five districts' English/language arts and mathematics standards; the CDE review forms of 190 districts' accountability plans (from the 1998–1999 Consolidated Applications); and a project review of 36 district accountability plans.

Major findings of the evaluation are summarized below.

The Status of Local Accountability Systems

Less than three years ago, the California Department of Education first directed California school districts to begin work on local standards-based accountability systems. Since that directive, some districts have made remarkable progress in designing systems, while others are still struggling over the first hurdles. Presented here is a summary of the study's findings about the status of local accountability system design and implementation in the spring of 1999, as well as about educators' conceptions of and views about accountability in general.

Most districts were still in the early stages of developing their local standards-based accountability systems.

A standards-based accountability system, a complex entity with multiple components such as content standards, aligned assessment and professional development, and consequences and incentives, takes time to plan and implement. A majority of districts (56.6 %) reported on the district survey that their standards-based accountability systems were still "in development." But some components enjoyed wide implementation.

The most common elements implemented to date are content standards and assessments. Over 57 percent of district survey respondents reported that their districts had "fully implemented" grade-level content standards in English/language arts and mathematics. The use of multiple assessment measures to determine whether students meet grade-level standards was similarly well under way in more than 60 percent of districts. For about 75 percent of these districts, however, implementation of content standards and multiple

assessment measures was still a relatively recent thing, having occurred only in the past two years.

Unlike content standards and assessment measures, interventions and rewards have rarely been implemented. Less than 30 percent of district survey respondents reported that their districts had fully implemented district intervention strategies for low-performing schools. About 85 percent of districts said that they had *not* implemented district rewards or incentives for high-performing or improving schools.

These data, taken as a whole, suggest that California school districts have been moving ahead in developing their accountability systems. Their focus, however, has been on the instructional framework — content standards and assessments — and not yet on the supporting structure (such as interventions for schools) to make the system work. Interviews generally confirmed this impression. For example, some district administrators who were interviewed suggested that although their districts have “assessment systems,” they do not yet have full-fledged accountability systems.

Development of accountability mechanisms is a long, sometimes messy process.

The process of developing a district accountability system is far from quick and straightforward. This may stem from districts’ efforts to use a community-based approach to accountability system development. Survey results and interviews suggest that districts have typically sought wide participation in the development of their accountability systems. In many districts, teachers, principals, other school administrators, schools board members, parents, and business and community members have been involved, along with district administrators. Although the inclusion of such a wide range of stakeholders in the development of the system may help engender support for it, the drawback is that the process takes time, especially when factoring in other matters, such as careful attention to state requirements.

The length of time that districts have been developing their accountability systems is highly variable, as is the amount of resources with which they have to work. Findings suggest that small rural districts and medium-sized districts have had greater difficulty in developing and implementing their systems than large districts, most likely due to limited capacity.

People hold varying views on what constitutes accountability, but many equate it with assessment.

Study results show that conceptions of accountability are often not uniform — even within the same district or school. In more than one district, different district administrators who were interviewed did not even agree about whether their district *had* an accountability system, *per se*. The same was often true at the school level. In one school, for example, the principal remarked, “The factors that have the most influence over teacher practice are test

scores and district accountability.” A teacher in this school, however, said, “There is none,” in response to a question about accountability.

Clearly, accountability means different things to different people. One of the more commonly held conceptions, however, is that accountability is synonymous with assessment. Even before the PSAA put heavy emphasis on the Standardized Testing and Reporting (STAR) test — the Stanford Achievement Test, 9th Edition, Form T (SAT-9), many district and school staff viewed the SAT-9 as the embodiment of accountability, as represented by the following comments from teachers who were interviewed:

We are held accountable for our student scores on the SAT-9. The district expects scores to go up.

The SAT-9 test is our accountability system.

We have the SAT-9. We seem to base everything on it.

Other important components, such as consequences and incentives, were mentioned far less frequently in response to general questions about accountability, both on the survey and in district and school interviews.

There is strong support for the concept of accountability — less so for the details of its implementation.

Overall study results suggest that most people, particularly at the district level, believe in the potential of accountability systems to raise student achievement. Many district survey respondents, for example, reacted positively to the PSAA, as shown by the following representative comments:

It's huge!...In the long run, it will be valuable and help our kids.

In the long run, it should create more sense of responsibility for the staff, since they are being asked to be accountable for what is happening in the classroom.

Mandated accountability will drive the district to review current programs and practices and make adjustments for improved student performance.

Nevertheless, many district administrators were concerned about some of the specifics of accountability system implementation, indicating that accountability still has a long way to go before it can meet the high expectations districts and schools have for it. The range of concerns is illustrated by the following comments:

In favor of accountability but needs to be aligned to curriculum standards.

Once the augmented test is improved, the pending legislation will be positive. It will drive change. ...However, the augmented test in its current configuration will not give us meaningful information.

We are concerned that important factors (poverty, transiency, LEP numbers) will not be taken into account.

Theoretically, it is sound and necessary. Timelines are unrealistic and the financial support is not there (i.e., increased need for extensive staff development while at the same time SBCP days were eliminated).

The state needs a “master plan” to help guide the school accountability process. State legislation is coming faster than districts can implement and assess the outcomes.

Content Standards

Content standards — what students should know and be able to do — are the first building block of a standards-based accountability system. If schools are going to be held accountable for ensuring that students master certain content, that content must be clearly specified. Content standards perform that function. In California, districts developed their own local standards, but these standards were required to be aligned with the standards adopted by the State Board of Education. This section summarizes the evaluation findings about district development and adoption of content standards, the relationship of the local standards to the state standards, and the impact of standards on curriculum and instruction in the classroom.

Almost all districts have only recently completed a lengthy, broad-based standards development process.

About 90 percent of district survey respondents reported that their districts have adopted mathematics and English/language arts content standards, although not all districts had fully implemented them. For most of these districts, this adoption has taken place in the past three years. In almost every district, all schools and students are held to the same set of standards.

“I believe standards are important. You have to know where you’re going before you take off or you’re just going to be everywhere. ... I think standards have also helped us talk about what we do.”

— Teacher

The development of standards appears to be a lengthy, complicated, iterative, and ongoing process. For example, in one district where committees of principals and teachers developed key indicators of what students should be able to do at each level and provided examples of student work, the development of standards (and aligned assessments) has been a four-year process of revisions. In fact, as of the time of the interviews in this district, it was still considered to be an ongoing work-in-progress.

Districts claim that the state standards were used to develop or modify local standards, but the rigor of local standards remains uncertain.

District survey results indicate that, after the state standards were adopted by the State Board of Education, most local districts either adopted the state standards in whole or used them for comparison purposes in the development of new or modification of existing local standards. Given that many districts already had developed local standards prior to the establishment of state standards, alignment of the existing local standards with the state standards was sometimes a source of frustration.

Over 85 percent of district survey respondents reported that their districts had either compared the district standards to the state standards for rigor or had simply adopted the state standards. Among districts who made the comparison, most respondents (90 percent for English/language arts and 80 percent for mathematics) reported that the district standards were “as rigorous” or “more rigorous” than the state standards. However, an independent analysis of some districts’ local standards, conducted by project staff, raised questions about such claims, revealing that districts’ appraisals of rigor may have been overly optimistic.

Local content standards have yet to make a consistent impact at the classroom level.

Many district survey respondents reported that the majority of teachers are using local standards in the classroom. For example, 54 of 108 respondents (50%) estimated that the district mathematics standards were reflected in the classroom teaching of 75 percent or more of district teachers. School-level observations and interviews, however, suggested that the classroom-level implementation of the standards may not be at the high level suggested by the district responses. There appeared to be a widening disparity in how content standards were understood and used as they filtered down to the school and classroom levels.

“I don’t think teachers are very tuned to the standards. There’s confusion. Our people are lost. Our standards aren’t exactly the same as the state’s and there’s confusion about why they would have different standards.”

—Principal

For example, at the school level, interviews revealed considerable confusion and frustration about different sets of standards or about having standards changing. One mathematics teacher remarked:

At all three levels [national, state, district] we have been bombarded. When we, as the math department, were given the standards, the NCTM, state, and local standards all conflicted with each other. We adopted the NCTM standards, which used to be closely aligned with the state standards. The state standards are what we are tested on. The new state standards are very different...It seems like a moving target. Every couple of years the state comes out with a different strategy and we all change and then things change again.

Moreover, interviewers found a great deal of variability in how standards were used and implemented, even within districts and within schools. A teacher in one school claimed that the district standards “are on the wall in every classroom” and said that “our jobs as teachers are linked to these standards.” However, another teacher in the same school commented, “As for the district standards, I’m a new teacher and not aware of what they are exactly.”

Perhaps as a result of the lack of consistency and the high level of confusion regarding standards, direct impact of the standards on curriculum and instruction appeared to be relatively low or somewhat superficial in most schools that were visited. Several mathematics teachers who were interviewed did not mention standards at all when asked how they decided what content to teach, and a few suggested that the district standards “did not apply” to them or to their students. Some felt that the state standards hold unrealistically high expectations for students and are overwhelming.

The low level of impact of standards on instruction may stem from insufficient alignment between curriculum and standards and from inadequate time spent on standards-related professional development. When asked to rate the extent to which district-adopted instructional materials are aligned with district content standards, less than half of survey respondents (for both mathematics and English/language arts) indicated a “great” level of alignment, although about 40 percent indicated a “moderate” level of alignment.

As for professional development, only about 25 percent of district survey respondents reported that, in a given year, they require more than three days of standards-related professional development for teachers.

“Our major job next year is to align curriculum, see if we’re achieving the standards, and understand what the assessments show about changes that need to be made.”

—Principal

“What they are doing is an excellent goal, it’s just too fast without enough training time for teachers.”

—Teacher

Nearly as many (roughly 20 percent) said that they require less than one day. For many teachers, focusing instruction on content standards would require a major shift not only in teaching practice but in teaching philosophy — a shift that would be unlikely to occur in the absence of sustained, meaningful professional development and capacity-building.

Assessment Measures

Along with content standards, assessment constitutes another key component of an effective accountability system. If students are to demonstrate their mastery of the content standards, they must have a forum through which to do so, and assessments that are aligned with the standards provide that forum.

The statewide Standards-Based Accountability System required districts to have at least one assessment measure in every grade and multiple measures for at least one grade level in each of three specified grade spans. In 1998, the state established statewide criteria for combining multiple measures and required that the SAT-9 be one of the measures used. The PSAA, which replaced the Standards-Based Accountability system, continued the requirement that almost all students be tested with the SAT-9. District-determined multiple measures, however, are no longer required, although the California Department of Education has encouraged districts to continue their use (Eastin, 1999). This section presents the major evaluation findings regarding districts' use of and views about various assessment measures.

The SAT-9 is pervasive, but its influence is viewed as highly problematic by many districts and schools.

Not surprisingly, almost every district included the SAT-9 as one of their assessment measures. Nearly all survey respondents, for example, reported that the SAT-9 is used at one or more grade levels for accountability purposes. Data from other sources supported the survey finding that most districts are indeed using the SAT-9.

As a mandatory, high-stakes test, the SAT-9 exerts considerable influence at the school level, affecting what professional development opportunities are offered and how classroom time is spent on test preparation. At many schools that were visited, the influence of the SAT-9 goes beyond preparing students for the test and extends into the realm of shaping the curriculum itself. Many people who were interviewed spoke of the SAT-9 “driving instruction.”

However, the SAT-9 is viewed by district- and school-level personnel as inherently flawed for the following reasons:

The SAT-9, in its current form, is not aligned with state standards. Many district and school staff cited this mismatch as a major flaw of a supposedly standards-based accountability system. As one district survey respondent wrote, “Standardized tests not aligned to standards cause problems in the implementation of the standards.”

“Student achievement is being gauged by a test NOT in alignment with State Standards.”

—District Survey Response

“I don’t really believe in teaching to a standardized test. I think that too much importance is placed on them. At the same time, I’m tugged in that direction, because everybody thinks it’s important. ... So, I have to honor it.”

—Teacher

The emphasis on the SAT-9 promotes “teaching to the test” — that is, limiting the curriculum to what is on the test. The SAT-9 plays a prominent role in state accountability policy. In fact, it is currently the sole indicator of school performance for the Academic Performance Index (API). As mentioned above, many people who were interviewed suggested that the SAT-9 is driving instruction; not everyone, however, views this as a good thing. For example, one mathematics teacher remarked:

I get the impression from state government that we need to teach to the test. I mean, who cares about content anymore in the math class? We teach to the test.

Similarly, a principal commented that, as teachers teach more and more narrowly to the test, important things may be getting left out of children’s education. The lack of alignment between the test and the standards exacerbates this problem: if teachers are teaching to the test and the test does not assess mastery of the standards, then teachers are not teaching to the standards. As one teacher put it:

I think the problem we have right now is that the test and the curriculum are based on different standards, and they haven’t brought them in line. And I’d like to see the test follow the curriculum — or, decide what the curriculum should be, establish the statewide standards, or national standards, or whatever the heck we’re going to use, and then make sure the test follows that. And not the other way around. I don’t want a curriculum chasing the test. I want the test to match the standards.

The “augmented” standards-based items included in the SAT-9 are considered overly difficult and unfair. The state, aware that the basic-skills SAT-9 test provided by the publisher is not aligned with the state standards, commissioned the development of new standards-based, “augmented” items. However, many of these items, which were used for the first time in the

spring of 1999, had technical validity and reliability problems. Moreover, many school-level staff reported in interviews that the augmented items were too difficult for students. One teacher, for example, commented that SAT-9 testing had been “a whole week of upset and tears” for her students, whom she said are among the best at her school. Several teachers did, however, indicate that they were planning to adjust their curriculum to cover the content of the augmented items, suggesting that inclusion of the new items may indeed be spurring teachers to teach to the standards.

The requirement that all students be tested in English is unfair. Several district survey respondents raised concerns about the requirement that all students, especially English language learners, take the SAT-9. People commented that this requirement yields misleading data about student, school, and district performance. As one survey respondent wrote:

Requirement to test all Limited English Proficient students in the SAT-9 yield invalid scores (student & school) and thus makes it difficult for us to compare schools since the student populations tested are so different.

Testing takes valuable time from instruction. Another area of concern with regard to assessment is the amount of class time needed to administer and prepare for tests. (In this respect, the SAT-9 is not the only culprit; other assessments also are responsible.) Several principals and teachers who were interviewed mentioned that too much time was being devoted to testing, and that testing and test preparation were cutting into instructional time.

Most districts use and value multiple measures, but fear they will be abandoned because of new state accountability requirements.

Consistent with state requirements that existed prior to the passage of the PSAA, most districts report using a wide array of multiple measures, including:

writing samples (86.3% of survey respondents);
grades (83.2% of survey respondents);
teacher evaluation/judgments (55% of survey respondents); and
district criterion-referenced tests (55.7% of survey respondents for English/language arts; 57.3% for mathematics).

Although the use of each of these types of measures may carry its own problems (such as reliability/validity determinations, the expense and time involved in developing local assessments, and the alignment of local measures with content standards), districts voiced strong support for the concept of multiple measures. Many districts even developed multiple measures for every grade level, going beyond the state requirements.

Indeed, districts' support for multiple measures lies behind one of their major criticisms of the centerpiece of the PSAA, the Academic Performance Index (API), which at present relies solely on SAT-9 scores to rank schools. Many district personnel expressed concerns that the state's current emphasis on the SAT-9 — which is not fully aligned to the state standards — as the sole indicator of student performance will lead to the demise of multiple measures and may drive the education system in harmful ways. Survey comments to this effect included:

CA swinging back to dependence on standardized test only rather than multiple measures will be devastating.

The new [accountability legislation]...basically does away with multiple measures. The SAT-9 is a snapshot in time whereas multiple measures is accountability system for the entire year.

Use of Data

Once students complete assessments, whether they are the SAT-9, district-developed assessments, or something else, the results must be analyzed to determine how well students have performed, not only on an individual level but also on classroom, school, subgroup, and district levels. Only when these analyses have occurred can judgments be made about the adequacy of performance at each level. Moreover, assessment results can be used to determine areas of strength and weakness in student performance, and their analysis thus can serve as one step in the process of providing curriculum and instruction to target the weaknesses. All of these analyses, however, require some sophistication in the use of data. Summarized here are the major evaluation findings regarding what data are looked at, by whom, and for what purposes.

Districts — perhaps because of capacity factors — report that they are more likely than schools and teachers to analyze and use student performance data.

According to survey results, district personnel make greater use of data than school level personnel. About 80 percent of survey respondents indicated that district accountability and assessment personnel examine and analyze student assessment data “to a great extent.” In contrast, only 33 percent of respondents reported that teachers in their district examine and analyze data “to a great extent.”

One explanation for this pattern is that there is greater capacity for data use and analysis at the district level than at the school or classroom level. More than 75 percent of survey respondents rated the capacity to analyze data at the district level as “good” or “very good”; only about 20 percent rated capacity to analyze data at the classroom level as “good” or

“My goal...is that all levels of the system, particularly teachers and administrators, understand the data they get that is provided centrally, know how to supplement it with meaningful data from the site, and know how to use that information on a regular basis to change instruction.”

—District Administrator

“very good.” The disaggregation of data by various factors (such as gender and socioeconomic status) is also more likely to occur at district levels than at school or classroom levels. Many districts do, however, recognize the data limitations at the school level and are trying to overcome them by offering school-level personnel professional development opportunities on how to

use and analyze data more effectively. Districts’ capacity-building efforts regarding use of data were supported both by interview and survey findings.

Small districts have inadequate resources to analyze and evaluate data and to assist schools and teachers in using data to improve programs.

District interviews suggested that district capacity to use and analyze data may be a function of the size of the district. Large districts seem to have the expertise and technology necessary to conduct sophisticated analyses of data, whereas the capacity of smaller districts appears more limited. On the survey, over 60 percent of districts cited needs for professional development in the use of data, better data-use technology, and more staff with evaluation or statistical background.

According to districts, engaging in analysis of student assessment data to meet state and federal reporting requirements occurs more frequently than analyzing data to improve instruction.

Survey results suggested that the satisfaction of reporting requirements constitutes the most prevalent purpose for district-level use of student assessment data. When presented with a number of possible reasons why a district might collect and analyze student assessment data and asked to rate the extent of data use for each one, districts emphasized “to satisfy state and federal reporting requirements” far more than any other listed purpose. Indeed, 96.9 percent of respondents indicated that their district uses data for this purpose either “to great extent” or “to moderate extent.”

Other reasons that were rated highly on the survey included the improvement of instruction (cited as “great” or “moderate” by 86.9 percent of respondents), the identification of students needing assistance (82.9%), and the gauging of student subgroup performance across the district (79.2%). Substantial use of data to identify teachers needing assistance was cited by only a few districts; 46.8 percent of districts said that they do not use student assessment data to identify teachers who need assistance *at all*. Similarly, 48.4 percent of districts said that they do not use assessment data to rate or rank district schools.

School-level interviews confirmed that teachers are less likely than principals and district administrators to use and analyze student assessment data.

While districts reported that they are using and analyzing data more than personnel at the school and classroom level, interviews with principals indicated that they are using and analyzing data to a significant, yet varying, degree. Several principals who were interviewed indicated that a substantial amount of data use takes place at the school level. However, teachers who were interviewed rarely mentioned using data at all. Thus it seems likely that data analysis has had little, if any, impact on instruction in most classrooms.

Consequences and Incentives

By some definitions, the very heart of an educational accountability system is consequences. The notion of educational “accountability” implies that someone, somewhere is being held accountable, or responsible, for student achievement and will face consequences if achievement does not rise. Thus, any discussion of accountability is incomplete if it does not include some mention of (a) performance targets for schools, the specific goals that are set for increased achievement, and (b) the consequences for reaching — or not reaching — these targets. This section summarizes the major study findings about what types of performance targets are set by districts and the consequences and incentives districts provide to reach targets.

Most districts established performance targets for schools, but their use and approach varied considerably by district.

Prior to the passage of the PSAA, districts were largely free to establish their own performance targets. Because Title I schools were required to do so, most did: 61.1 percent of district survey respondents reported having Adequate Yearly Progress (AYP) goals for Title I schools. The numbers of non-Title I schools with targets were, unsurprisingly, lower. In fact, over 30 percent of respondents said that their districts set no performance targets for non-Title I schools.

Only a few districts offered rewards for good performance, and most districts were only at the beginning stages of thinking through consequences for low school performance.

On the district survey, only 5.3 percent of respondents said that their districts offered incentives or rewards to individual schools for high student performance or for improving performance. As for consequences for low-performing schools, interviews suggested that

“If we really believe we are a profession that can deliver the goods on student achievement, then we should be willing to accept rewards for success and consequences for failure. We should adopt a no-excuses philosophy.... We must be willing to accept responsibility regardless of the background of our students.”

—District Administrator

districts that were instituting such consequences were still in the very early stages of doing so, and there appeared to be different perspectives about the scope and implementation of consequences within and among these districts.

On the survey, districts were more likely to report using some form of supportive intervention for schools than they were to report implementing sanctions. The most frequently cited

interventions included the assigning of experts to help schools develop improvement plans, the tailoring of interventions to individual schools, and the provision of professional development opportunities to schools. Only 1 respondent out of 127 indicated school reconstitution as an important district policy for schools with low performance.

Principals are more likely than teachers to be held accountable for performance at the school level.

State policymaker and district interviews both pointed to the principal as the primary person at the school level whose job was on the line for student performance. In many cases, however, the accountability of the principal remained somewhat informal or even hypothetical. In nearly every district where interviews occurred, administrators reported that principals could, in theory, be removed from their positions for accountability reasons *per se*, but none actually had been removed as of the time of the interviews.

Holding teachers accountable was viewed by most people as legally problematic or as too complicated. Teachers who were interviewed were also unclear about whether they were subject to any consequences for low student performance. Even within districts, teachers held differing perceptions about the extent to which they would be held accountable. Several said that they did not know whether they were; others suggested that there were no consequences tied to accountability. As one teacher commented, “Nothing is being done.... There’s a greater awareness of what the [test] scores mean, but as far as accountability at a personal level, I don’t think there is any.”

Two other groups that might be held accountable for academic performance are students themselves and parents. Interviews indicated that students are being held increasingly accountable for their performance through policies prohibiting social promotion or requiring summer school. Parents can be held accountable for student attendance, but not for student achievement. Some districts were experimenting with ways to increase parent involvement.

State policymakers and district administrators who were interviewed acknowledged that, if schools, teachers, and students are going to be held accountable for reaching certain goals, they must be provided with the levels of support necessary to achieve those goals. One district administrator, for example, remarked, “The notion that we’re going to punish kids, and hold them over, when they haven’t had access to quality instruction, isn’t right.”

“The pressure to improve test scores has increased while teacher support has not, [leading to] increased anxiety level and low morale.”

—District Survey Response

When given the opportunity, districts raised many concerns about the state’s new Academic Performance Index ranking system.

The PSAA introduced a statewide system of ranking schools (the API) and of providing interventions and incentives for schools (the II/USP and the HA/ISP). On the district survey, some respondents indicated a belief that the ranking would be counterproductive, “defeating the goal to improve,” as one person put it. Several respondents did not object to ranking schools *per se* but felt that only schools with similar student populations should be compared. “No problem being accountable,” wrote one respondent, “Just don’t compare schools of high and low SES and high and low LEP counts.”

Another concern expressed by some districts about the new legislation is that sanctions involving a change in school staff may exacerbate already-difficult conditions for the lowest performing schools. As one survey respondent said, “I agree with the intent but the playing field is not level. We anticipate changes in administration at schools with high ELL populations due to schools not showing necessary growth. Thus, schools most needing stability will likely go through greatest change.”

“I have grave concerns about sanctions for low-performing schools. ... Sanctions will encourage the dedicated experienced teachers to transfer to higher-performing (higher socioeconomic) schools.”

—District Survey Response

Lastly, many districts felt that the goal of five percent growth per year was unrealistic and unfair. A couple of district administrators who were interviewed remarked that achievement does not improve in a constant, linear fashion, but rather occurs in spurts. Schools with achievement that is already high were particularly concerned about showing five percent growth each year.

Impact of Standards-Based Accountability Systems

Because most districts' standards-based accountability systems have been in place for less than two years, it is too early to know for sure what effects, if any, these systems are having on student achievement. Even where evidence of improvement — or, in a few cases, decline — exists, so many different reforms and initiatives have been undertaken simultaneously that one can only speculate about whether the observed changes are attributable to accountability measures.

Although improved student achievement is the most important goal, other outcomes, such as teacher morale, also are important because ultimately they affect student achievement. For these types of outcomes, it may not be too early to expect effects or to infer causality. Presented here is a summary of the study's main findings about the effects of accountability on factors such as policy, practice, achievement, and teacher morale.

Districts and schools report that accountability has had positive effects, especially on curriculum, instruction, and assessment practices.

Both survey results and interviews reveal that district and schools express a cautious optimism about the positive impact of accountability. Representative survey comments about the positive effects of accountability included:

Has everyone focused on student learnings and student outcomes.

Accountability has brought increased focus on achievement.

Accountability has positively impacted how we teach and assess.

As these comments suggest, many believe that greater accountability has led to a stronger focus on student learning. Over 80 percent of survey respondents reported that accountability has had a positive effect on district-level curriculum and instruction policies and on district-level assessments; 88.6 percent reported a positive effect on curriculum and instruction practices in schools. Over 70 percent of districts surveyed also said that accountability had had a positive effect on classroom assessments.

Moreover, in the opinion of 42.8 percent of survey respondents, their district accountability system or particular accountability components had exercised a positive effect on standardized test scores. Even more powerfully, 64.1 percent of respondents postulated a positive effect on student achievement as measured by other indicators of academic performance.

At the same time, districts and schools report that accountability has lowered teacher morale.

Although most districts that were surveyed reported a positive effect of accountability on most educational areas, over 40 percent of respondents reported that teacher morale had been adversely affected. Many survey respondents commented that teachers are feeling considerable stress and frustration as a result of accountability requirements. Representative comments included:

Teachers are anxious about how they will be evaluated and possible job termination if their students' scores are low or do not show growth.

Teachers are feeling tremendous pressure to succeed with every student even though many factors of student performance are not in their control.

Teachers are feeling overwhelmed by the number of changes which have contributed to some expressions of frustration and inability to feel they are keeping up with expectations.

Some school-level staff — principals and teachers — echoed these concerns in interviews. “I am really frustrated by it all,” said one teacher.

However, accountability does not necessarily affect all teachers negatively. Some interviewed teachers did make positive remarks about accountability and even about being held accountable, and several district survey respondents also commented that accountability had exercised a positive effect on teachers or on classroom instruction. Such survey comments included:

Teachers always enjoy meeting to set directions in their subject area. Assessing these areas of direction gives them feedback that they haven't had in the past — they like knowing how their efforts are paying off.

Teachers understand the target and can now make sure students are prepared.

Accountability is a role that the teachers and administration welcome — creating a positive direction for all staff, parents, and students.

Challenges and Assistance

The development and implementation of an accountability system is by no means a simple, straightforward task. Many challenges present themselves along the way. For example, not only must the system be designed in ways that are likely help raise student achievement —

a difficult charge even under the best of circumstances — but it must be designed and implemented within the constraints of available time and resources. The nascent system must also be coordinated with other existing or emerging policies, programs, and initiatives. Moreover, the effectiveness of the system will rely on broad support and buy-in from multiple stakeholders. This section summarizes districts’ comments about these challenges and others they have faced, as well as what they have found and would find helpful in responding to the challenges they face.

Many district officials expressed frustration and confusion about the lack of a consistent, coherent, and fair statewide reform agenda.

In interviews and survey comments, district administrators expressed great concern about their ability to keep up with the rapidly changing set of reforms initiated at the state level. The addition of new requirements without the elimination of older ones causes a strain on districts, and it also sends mixed messages about what is important. Further, district and school personnel are frustrated by the frequent changes in state requirements and the lack of coordination between different state initiatives. State level “practices” that survey respondents mentioned as having hindered their district in accountability development included the following:

Lack of consistent requirements—we need a sustained focus over enough time to yield meaningful information.

Inability of state to evaluate and adopt a system that is consistent, based on proven implementation, with professional development attached.

Constantly changing state legislation.

State system not consistent year to year.

Moreover, nearly 60 percent of district survey respondents said that they do not believe that the expectations and requirements for districts with regard to accountability are fair and reasonable. The sources of “unfairness” cited by many districts included the use of a norm-referenced test to measure mastery of content standards, lack of consistency and coordination at the state level, unrealistic state expectations for student achievement, and inattention to issues of equity.

Districts have faced many specific challenges in developing and implementing key components of an accountability system.

Over 50 percent of survey respondents indicated each of the following as one of their district’s top five challenges in developing and implementing an accountability system:

aligning curriculum, instruction, assessment, and content standards
dealing with limited resources (time, staff, financial support)
finding or developing valid and reliable forms of assessment

Ensuring equity of the system for all schools and populations subgroups was another frequently noted challenge on the survey, as was implementing professional development for teachers in the use of content standards. District interviews suggested that yet another key challenge is gaining buy-in from teachers and others at the school level so that reforms are actually implemented. Districts also indicated both on the survey and in interviews that communication-related issues, such as lack of clarity in state directives and delayed information, have created challenges for them. Limited resources, too, appear to have been an issue with many districts — and at the state level.

The challenges presented by the PSAA more or less parallel the more general concerns held by districts. For example, many districts raised questions about the time and resources needed to implement the new policy (especially in conjunction with other requirements) and about the level of coordination between the PSAA and other initiatives. Representative survey comments included:

Having one accountability/assessment package that is coordinated is much preferable to a piecemeal approach. We are concerned that the popularity of this hot initiative (assessment/accountability) will result in a flurry of sometimes conflicting expectations...not supported with resources, and nothing else has gone away, so it has added another layer of tasks.

It seems to be coming down to the local level at a fast and furious pace. It is difficult to thoughtfully implement new policies and programs and accountability measures given quick turnaround time.

Although most districts indicated that they would make a good faith effort to abide by the new system, many indicated that they need time and resources to determine what strategies are best to remedy problems identified by accountability measures.

Districts have received assistance from a wide range of organizations in the development of their standards-based accountability systems.

Districts turned to a variety of organizations for help in implementing their standards-based accountability systems. Many districts (63.2%) reported that they have received assistance from the California Department of Education in developing, implementing, and refining their accountability systems. Other frequently cited sources for support included county offices of education (marked by 69.2% of survey respondents), other districts (51.9%), professional consultants (40.6%), and the Statewide System of School Support (32.3%). Districts were particularly positive about the assistance and documents they received related to content standards and multiple measures.

Districts would like further assistance related to accountability.

When asked, “Would you or others from your district like to have an opportunity for more professional development related to standards-based accountability?” over 80 percent of survey respondents answered in the affirmative. The types of professional development desired were quite varied; areas that were identified particularly often included the development of assessments and the alignment of assessments, standards, curriculum and instruction. Another area in which many districts requested assistance was in the analysis and use of student assessment data; the phrases “analyzing data” and “data analysis” appeared again and again in districts’ written survey comments about what types of professional development they would find helpful.

Improvement of school and student performance and “putting the pieces together” were other areas in which districts requested assistance, perhaps through networking with other districts. Districts would like concrete, specific, proven strategies, especially ones to use with low-performing schools and students, as indicated by the following survey comments:

More workshops highlighting best practices.

SOLID RESEARCH ON WHAT REALLY WORKS for improvement.

Models of systems employed by districts of a similar size and with limited district office staff.

Putting-all-the-pieces together ideas from successful systems.

Increased resources and greater stability and consistency of state policy also appeared on the “wish list” of many districts.

Implications and Recommendations

Overall, the study revealed a high degree of buy-in and commitment to the concept of accountability up and down the educational ladder in California. District administrators, principals, and teachers know and accept the need to be accountable for improving student achievement. Nevertheless, many districts have experienced considerable frustration over various aspects of the implementation of the state accountability initiatives. This section outlines the general and specific recommendations that emerge from the study’s findings.

For accountability to be effective, it must be systemic and coherent in nature, and it must be clearly communicated at all levels of the educational system.

It is not enough for individual pieces of accountability to be present — they must reinforce and align with one another. Much more work is needed in California before such an accountability system is complete and leads to improved student performance in a way that is consistent with state academic standards.

To have such an effect, state content standards must be aligned with local standards, which must in turn be aligned with performance standards. Curriculum frameworks, assessments, instructional materials, professional development, and even teacher preparation programs all must be consistent with the standards. These and other components must be appropriately aligned throughout the system — all the way from the policymakers in Sacramento to each classroom teacher. In California, however, many of these elements have not as yet been fully developed, nor are they coordinated with one another. Furthermore, the loose coupling between state policy and local implementation by teachers poses a significant challenge to the effectiveness of the accountability system.

Clear communication and consistent messages about accountability at all levels of the educational system — from the state to the district to the school to the classroom — would also increase the positive impact of accountability on student performance. The public and those at the local level would benefit from a clearer understanding of what accountability is, how it is supposed to work, what its components are, and how the components are interrelated. In particular, clear communications about what is expected must reach the level of the classroom for accountability to help raise student achievement.

In addition, greater stability in policy at the state level would facilitate the implementation of accountability at the local level. If policy changes are necessary, they should be based on evidence collected from those who are most affected, namely local administrators, educators, and students. California has frequently changed direction, adopting one reform after the next without allowing sufficient time to fully implement or evaluate any — much to the frustration of district personnel. If California is to have an accountability system that enhances student achievement in the right direction, it must continue to strive for a standards-based, stable, and consistent approach, and send clear signals to districts and schools.

For the most part, California should “stay the course” with developing the existing components of its accountability infrastructure: standards, assessment, and a system of interventions, rewards, and sanctions. However, no approach is perfect from the start. Modifications may be necessary to rectify unintended consequences and ensure the system is meeting its primary objective. The caveat is that *any changes should be informed by systematically collected evidence from the local level about what is working or not working and why.*

1. *Alignment Inventory*

An outside independent group should conduct a periodic “alignment inventory” of current state education policies. This group should examine the status of development and the degree of coherence among the several legislatively-mandated components of the current accountability system. This review would include a *horizontal* analysis of key state policies, such as PSAA (SB X1), STAR (AB 265), High School Exit Exam (SBX 2), and the various other state curriculum and instruction and professional development initiatives. A *vertical* analysis, examining how well such policies are communicated and implemented from the state, to districts, to schools, to the classroom level should also be part of this inventory. The appointment of the outside body, which should report its findings on an ongoing, regular basis to CDE and to the State Board of Education, the Legislature, and the Governor, should be one of the first charges to the Accountability unit within CDE.

2. *Accountability Evaluation*

The Governor and the Legislature should adequately fund the evaluation currently mandated by the PSAA in order to provide a comprehensive, rigorous look at the effects of the new accountability program. This evaluation of the PSAA would serve as an important source of information on how well accountability policies are understood, implemented, and used to facilitate change in classroom teaching and to improve student outcomes in desirable directions. Adequately funded, the evaluation could help ensure the availability of rigorous and generalizable information to inform policymakers about any modifications necessary to the existing accountability system. The evaluation should pay close attention to matters of equity, such as the impact of the accountability system on student subgroup populations and schools in a wide range of settings.

3. *Use of the World Wide Web*

CDE and the State Board should continue to ramp up their use of the World Wide Web in communicating accountability policy to all stakeholders within the system, from district personnel to teachers to parents to the general public. Since the enactment of the PSAA, the Web has served as an invaluable communication resource. Further development and resources should be spent on updating and making more comprehensive the PSAA and STAR portions of the CDE Web site. The “alignment inventory” could also be housed on the CDE Web site as a way of providing local districts and educators with the “big picture” perspective on the status, interrelationship, and alignment of various components of the state’s accountability, assessment, and standards policies. In addition, some state and district resources should be focused on ensuring that teachers are familiar with the Web site, have easy access to it, and use it routinely.

Assessments for accountability purposes should measure student progress toward content standards; if they do not, they will divert attention away from content standards.

Ideally, in a completely articulated standards-based accountability system, schools should focus on teaching curriculum based on content standards, and assessments should merely be a measure of how well students have mastered the standards. However, the findings of this evaluation clearly demonstrate that, as a mandatory, high-stakes test, the SAT-9 often drives curriculum and instruction practices to a much greater extent than the content standards *per se*.

This is not surprising, given the state's emphasis on assessment. And were the SAT-9 fully aligned with the content standards, the emphasis on the test might help bring about the desired effect of student mastery of the standards. Yet there are serious concerns about the extent to which the SAT-9, even with its augmented sections, is aligned with the standards. Even should scores on the SAT-9 rise significantly over the next few years, an important question would be whether the rise is truly indicative of the type of student improvement desired.

Recommendations on Assessment for Accountability

4. Standards-Based Assessments

If content and performance standards are to be the drivers of the accountability system, standards-based assessments must be developed as quickly and carefully as possible. It is essential that assessments used for accountability purposes, whether the augmented SAT-9 or a matrix test, reflect the state standards. The Governor and the Legislature should provide adequate resources to develop such standards-based assessments as quickly as possible without jeopardizing their validity and reliability. The development of the standards-based assessments should include the input of educators, and, when completed, the assessments should be carefully field-tested.

5. Inclusion of Standards-Based Assessments in the Academic Performance Index

As soon as valid and reliable standards-based assessments are available, the Academic Performance Index (API) should include them. As the development of the API proceeds, the addition of other valid and reliable measures will help ensure that schools are not being ranked based only on one, narrow indicator of performance — at present, the basic-skills test items of the SAT-9. As with other accountability components, the critical variable is resources; to do this quickly without compromising quality will be expensive. Cutting corners to limit the amount of money spent or extending the timeline risks having a state accountability system that is driven by an

indicator of performance that is not consistent with state standards or a comprehensive definition of school performance.

6. *Statewide Student Information System*

The Governor and the Legislature should expedite and fully fund the development of the California School Information Services (CSIS) to facilitate the inclusion of reliable comprehensive measures, such as attendance and graduation rates, in the API. Such a statewide student information system would benefit all districts, but smaller districts in particular, by providing the infrastructure necessary to collect reliable data on attendance and graduation rates, track students who change schools, and so on. Although expensive initially, it is likely that a statewide data system would actually be more cost-effective than the current approach, in which individual districts struggle to design their own systems. In the short term, as the system is being developed, the state should provide additional support, such as an optional data service funded through a combination of state and local contributions, to small and medium-sized districts.

Accountability measures should be accompanied by capacity-building activities, such as professional development for teachers on teaching to content standards and for teachers and administrators on using data to improve instruction.

While a set of strong, standards-aligned assessments would go a long way toward improving instruction and student performance in the desired directions, they would be bolstered by additional capacity-building activities. In particular, professional development around how to teach to content standards and how to use assessment data to inform classroom practices would strengthen accountability efforts.

While most districts do offer some professional development on content standards for teachers, the current offerings appear limited. Current state-level efforts to link professional development to content standards are clearly needed, and their continued existence should be encouraged. Additional support and resources for this professional development would be invaluable in moving California toward high standards for all students.

Another area requiring significant capacity-building attention is the analysis and use of data. Professional development around the use of data for people at every level of the system — district administrators, school administrators, and teachers — is necessary for accountability to have its maximal impact. The use of data to improve instruction — and ultimately improve student achievement — is a resource of tremendous potential value that, as yet, has largely been untapped. Moreover, the new, complex formulas used for the API and the implications of using test results to determine rewards and sanctions will require

unprecedented levels of understanding among state-, district-, and school-level personnel if they are to appropriately employ the formulas and interpret their implications.

Recommendations on Capacity Building for Accountability

7. *Professional Development for Teachers in Content Standards*

Local school districts should ensure that professional development programs are aimed at building teacher knowledge and skills related to content standards. These programs should be adequately funded and sustained over time; district priorities for professional development might need to be examined. In addition, teachers must have adequate time to participate in these opportunities.

8. *Capacity Building for Accountability-Related Data Analysis*

The Governor and the Legislature should fund capacity-building opportunities for teachers and administrators to learn about analyzing data to improve student achievement and school performance. CDE should serve as a broker for these services and provide statewide uniformity and quality control for them. The county offices and superintendents should serve as regional resources for this training.

9. *Clearinghouse of Exemplary Practices*

The Governor and the Legislature should fund the development of a Clearinghouse of exemplary accountability practices developed by districts and schools for raising student achievement. The Clearinghouse, which would be accessible through the Web, would feature information about practices from districts that have shown a successful grasp of implementing accountability at the local level. Through the Clearinghouse, districts and schools would be able to find out about “best practices” that were successful in settings and with student populations similar to their own.

10. *Inclusion of Accountability-Related Topics in Teacher Preparation Programs*

The California State University and University of California teacher preparation programs should specifically address issues related to accountability. These issues include teaching to content standards and using data from accountability mechanisms to improve instruction.

Conclusion

At its current stage of development, the concept of standards-based accountability enjoys considerable support in California as a reform strategy. However, the sense of optimism about its effectiveness is jarred by the reality of implementation. Accountability has proven to be a complicated task and that complexity is exacerbated by the compressed time frame in which it has been implemented. As one district Superintendent put it, “We are training for a marathon and asked to do 100-yard sprints!” The early timing of this evaluation in the long-term development process cannot be emphasized enough. It will be several years before a fully developed accountability system can be expected to have widespread, positive impact.

The recommendations generated from this evaluation are few in number, but all are important. They have implications and responsibilities for all the actors in education in California: the State Board of Education, the Governor, the Legislature, local districts, county offices of education, superintendents, and the CSU and UC teacher education programs. Some of the activities mentioned in the recommendations are already underway, but accountability brings a heightened emphasis to them.

The overarching recommendation from this report is the following:

Step Back, Review, and Align

Political leaders and educators need to step back, review where the California educational system is, align what already exists, and resist the temptation to move forward with new initiatives.

The authors of this report believe that attention to the issues raised by this evaluation will help bring about a coherent, cohesive accountability system that can help improve the achievement of all of California’s children.