

ACCOUNTABILITY AND ASSESSMENT

Accountability has become one of the most important issues in education today. Schools traditionally have been expected to account for their fiscal expenditures, and for their overall performance. Today's schools, however, are expected to be accountable for the performance of each and every student. This new level of close scrutiny on student performance is especially evident in California.

Today's schools are held accountable for student performance under two assessment systems: 1) the federal system of Adequate Yearly Progress (AYP); and 2) the state system, which uses the Academic Performance Index (API).¹⁰ Both systems use STAR (Standardized Testing and Reporting) scores and the California High School Exit Exam to measure school performance.

AYP: Under the federal No Child Left Behind Act (NCLB), 100 percent of students are expected to be proficient in reading and math by 2013-14. "Proficient" in California means: 1) elementary and middle school students scoring "proficient" on California Standards Tests in English language arts and math; 2) for high schools, 10th graders scoring 77 percent in English and 69 percent in math on the Exit Exam. The state has set annual benchmarks, called annual measurable objectives for the percentage of students who should be proficient in order for schools to make "adequate yearly progress" toward the 100 percent proficiency goal. All subgroups of students based on ethnicity, poverty level, disabilities and status as English learners must achieve these annual measurable objectives. In 2003 schools must also have tested 95 percent of students in each subgroup, had an API score of at least 560 or increased by one point over the previous year, and for high schools, achieved a graduation rate of at least 84.8 percent or shown improvement.

In 2003, 76 percent of elementary schools met these benchmarks, but only 42 percent of middle schools did so. As for AYP, 67 percent of elementary schools, but only 35 percent of middle schools and a mere 28 percent of high schools met the proposed guidelines. (Data from the California Department of Education, see EdSource, cited above.)

API: The API is a single-number indicator of the performance of a school's students on state tests. Based on their API scores, schools at each grade level are divided into 10 performance levels or deciles, and then ranked from 1 (lowest) to 10 (highest). In January each school receives a "base API" between 200 and 1000 based on STAR test results the prior year and also a growth target. Scores are calculated for the entire school and also for significant subgroups based on ethnicity and economic status. Schools with API scores below the state's performance target of 800 are expected to progress each year by 5 percent of the difference between their base API and 800. Significant subgroups' improvement goals are 80 percent of their school's growth target. For 2003, 82 percent of elementary schools met their growth targets, compared to 69 percent of middle schools and 67 percent of high schools.

Impact of AYP on API: The federal NCLB program has impacted California's accountability measure with a new emphasis on an intervention called "Program Improvement." However, several programs in existence before the federal program have continued: 1) The Immediate Intervention/Underperforming Schools Program (II/USP) of 1999 targets schools with scores in the bottom half of the API; 2) The High Priority Schools Grant Program (HPSGP), a supplement started in 2001 shifted the focus to schools with

¹⁰ See EdSource, Accountability Data: AYP/API, Card 21, 2004 Resource Cards on California Schools

scores in the lowest 10 percent of the API rankings. Under the NCLB program, only schools that receive federal Title I funds are entitled to receive program improvements. If schools fail to make adequate yearly progress for two years in a row, they become eligible for program improvement. In year one, the districts must inform parents of the school's score and must strengthen academic subjects, offer teachers professional development, promote parental involvement and allow students to transfer (with free transportation) to another school. In year two, the district must do all of the above plus provide supplemental services such as tutoring or after-school enrichment activities.

Corrective action begins in year three. At that time the district must do one or more of the following: replace appropriate staff, implement new curriculum, decrease the management authority of the principal, appoint an outside expert, extend the school day or year, or restructure the school. In subsequent years, the improvements include developing alternative approaches to school governance, including arrangements for outside intervention. Schools that make AYP stay at the same designated year of the program, and after two years in a row, are released from program improvement. In 2003, 10 percent of all California schools (5,467 schools) were in program improvement and another 4 percent (340) were facing corrective action.¹¹

California High School Exit Exam: To graduate, California requires students to pass a minimum number of courses. Schools districts, however, can require more than the minimum. State requirements include three (3) years of English, two (2) years of math (including Algebra I), three (3) years of social studies (including a semester in American government and civics), two (2) years of science, one (1) year in visual or performing arts or foreign language, and two (2) years of physical education. However, starting with the class of 2006, public high school students must also pass an exit exam to receive a diploma. (The original requirement was to cover the class of 2004, but the state Board of Education moved the date to accommodate public concerns that there was not enough time for schools to comply.) The test will be based on California's academic content standards. Starting in the 10th year of high school, students can take the exam, and if they fail, can retake it as needed before graduation. Accommodations are made for students with disabilities.

Are California high schools ready? This question is the subject of a study by The Education Trust-West which looked at graduation rates, the courses students take and how well students master the skills they need. Regarding graduation, although the reported rate is 87 percent, according to the report, only 70 percent of California ninth graders graduate four years later. Latino, African-American, and low income students complete high school at even lower rates (59 percent for African Americans, 57 percent for Latino students, compared to 81 percent for white students and 90 percent for Asians). (See "A Special Report: Are California High Schools Ready for the 21st Century?" The Education Trust-West, 2004, page 2.) According to the report, these results are an improvement over the year 1998, before California adopted the Public Schools Accountability Act (PSAA), when the rate of high school graduation statewide was only 63 percent.¹²

The California High School Exit Exam requires students to answer only 55 percent of the math questions correctly to earn a passing score. Moreover, the skills tested are sixth-to-eighth-grade skills. Yet only slightly more than half of the 10th graders taking the test for the first time in 2003 passed the math section. In English, only 67 percent passed. The scores of Latino and African-American students were even lower

¹¹ EdSource and California Department of Education data

¹² Ibid.

(35 percent for Latino and 31 percent for African American students in math). If the state had not postponed the exit exam, up to one-third of the class of 2004 could have failed to earn a diploma on the basis of the exit exam.¹³

Too few high schools make sure students take and succeed in the courses needed to pass the exit exam or to qualify for higher education. About half the students enrolled in classes of algebra II and geometry are performing below basic levels on the California Standards Tests, as are about a third of the students enrolled in chemistry and grade-level English.¹⁴ By one measure, some 49 percent of students who take the right courses and receive grades qualifying them to enter the California State University system need remediation in English.¹⁵

Recommendations to improve overall achievement and close the gaps include (1) the recruitment and training of qualified teachers at the high school level; (2) not only offering but requiring the A-G curriculum sequence (15 courses that include four years of English, three years of math, two years of science and two years of history/social science) be taken by all students; and (3) providing schools with the materials and supportive services needed for all students to succeed (aligning test books to state standards, providing model lessons and benchmark assessments aligned to California standards, offering intervention and more focused assistance to students who fall behind).¹⁶

Statewide Testing Program (STAR): All California students in grades 2-11 participate in the Standardized Testing and Reporting program (STAR) each year. Parents and schools receive individual student scores. Results for schools, districts, counties and the state are made public and posted on the Internet. Based on their student test results, schools are given an API score and ranked. The test results are also used to determine whether schools have made AYP under the federal program. STAR consists of California Standards Tests based on academic content standards—what students are supposed to learn. Also included is a norm-referenced test of basic skills, in which each student’s scores indicate performance relative to a national sample. For native Spanish speakers in grades 2-11 during their first year in public school, results are compared to grade-level peers in the 12 states with the highest number of Spanish-speaking students. Starting in 2003, the norm-referenced test was the California Achievement Test, 6th Edition (CAT/6), replacing the Stanford Achievement Test, 9th Edition.

The California Standards Test covers content in four subject areas: English language arts (reading, vocabulary and language arts for grades 2-11; fourth and seventh graders also take a writing test); mathematics; history/social science; and science (for high school students in specific subjects, such as biology). Scores are graded as basic, below basic, proficient and advanced. The goal is for all students to score at a proficient or advanced level. Students with disabilities who are unable to participate in the STAR program, even with accommodations, are tested using teacher observations and student records on tasks that are the building blocks of California’s academic content standards.

¹³ Op. cit., p. 7

¹⁴ Ibid.

¹⁵ Ibid.

¹⁶ Op. cit., p. 10-11

In 2002 for grades 2-11, no more than 40 percent of students achieved the desired results of proficient or advanced in English language arts. Scores for mathematics were somewhat higher in grades 2-4, but declined substantially after that. Comparing California students to a national sample of students, using the CAT/6, a score of 50 would place California students on a par with the national sample; a score below 50 means California students perform below average. In 2003, California students received scores of 50 or above only in grades 2-4 in math and spelling. In all other grades and subject areas, California students scored below the averages of a comparable national sample.

LAO Analysis of Current System of Accountability for Student Performance

An analysis of California's current system of accountability for student performance by the Legislative Analyst's Office (*Analysis of the 2003-04 Budget Bill*, Education section) sheds an interesting light on the impact of these assessment programs and on their effectiveness in improving student performance.

In a brief history of assessments, the LAO lists a series of programs that California has adopted and then discarded. The California Assessment Program (CAP) tested average student scores in mathematics, reading, writing, social science and science in grades 3, 6, 8 and 12. The results were designed to provide a reliable gauge of average achievement at the school or district level. Students took only parts of each test; the state aggregated results at the school or district level. This program was superseded in 1991 by the California Learning and Assessment System Focused on Performance Testing (CLAS), which tested students in reading, writing and mathematics in grades 4, 8 and 10. These tests provided individual student scores. The program was discontinued after questions were raised regarding its reliability. In 1995, the California Assessment Academic Achievement Act was initiated, setting the stage for the development and adoption of academic content and performance standards to be used for assessment purposes. Content standards for skills and knowledge students should acquire in a given grade were completed and adopted for (1) English language arts and mathematics in December 1977, (2) English Language Development (ELD) July 1999, (3) history/social science and science in October 1998, and (4) visual and performance arts in January 2001. (LAO, *Analysis of the 2003-04 Budget Bill*, Education Appendix 2)

Currently there are multiple state and federal assessment programs as well as federal intervention and sanctions programs for struggling schools. The Legislative Analyst describes this as a situation "creating administrative complexities that subtract from the focus on school improvement." (Ibid., p. E-114) Specifically, the state system (1) rewards schools for academic improvement, (2) provides external intervention for low-performing schools, and (3) sanctions schools that continue to fail after receiving external assistance. The federal system has several fundamental differences including (1) growth targets which measure different goals than those set by the state, (2) different entities responsible for intervention, and (3) sanctions and interventions which differ significantly from those of the state. (Ibid., p.E-119)

The federal system focuses on the percentage of students meeting proficiency targets, while the state's system focuses on growth in overall school achievement from year to year. As the proficiency standards are increased annually, per federal requirements, more and more schools will be unable to achieve these targets. "In our view, virtually every Title I school will likely be in Program Improvement within a few years."(Ibid., p. E-120) In contrast, the statewide API performance growth target is generally 5 percent of the difference between the school's score in a given year and the target of 800. There is also a target for significant ethnic and economically disadvantaged subgroups, which is 80 percent of the schoolwide target.

State and school district roles also differ: The federal system makes school districts primarily responsible for implementing interventions and sanctions, whereas the state not only is charged with intervening in

districts that are not meeting AYP, but must provide technical assistance to schools through a “statewide system of school support.”

To comply with federal requirements, California has put in place a dual system that keeps the current API system while also reporting AYP to the federal government. Schools receive a report of both scores. Currently, about half of the schools statewide make their API targets, but many schools are also identified for Program Improvement under the federal system. This sends conflicting messages to schools. A matrix created by the state classifying schools based on a combination of API and AYP “adds yet another layer of complexity to the system.” (Ibid., p. E-122)

The Legislative Analyst recommends that the Legislature seek to integrate the state and federal systems “to create clearer expectations and to focus resources on the neediest schools.” To achieve this goal, the Analyst recommends that a “restructured accountability system focus on providing technical assistance at the school district level to build capacity for districts to intervene” in needy schools. (Ibid, p.E-125) The LAO also recommends restructuring of the High Priority Schools Grant Program to serve as the primary assessment program for both state and federal purposes. This would require (1) making planning requirements uniform so that schools need submit only one plan for state and federal interventions; (2) modifying sanctions to coincide with those required by the federal law for corrective action; and (3) providing all program schools with funding for a specified number of years, not using significant growth as a trigger.(Ibid., p. E-127)

In its analysis of the federal program, EdSource notes that the timing “is inconvenient at best and in some cases very challenging.” The state is already committed to a whole set of standards-based reforms that are just becoming familiar. Further, budget cuts have left schools struggling. (EdSource, “No Child Left Behind in California?” Executive Summary, 2004)

Critical Perspectives Regarding Accountability and Assessment

Several scholars have argued that California’s current assessment programs fail to address a fundamental principle of educational accountability: “Responsible accountability systems must attend to whether students have adequate and equitable opportunities for learning.”¹⁷ “Although monitoring what students learn is certainly important, it cannot be the sole element of a responsible accountability system. To be responsible, an accountability system must also attend to whether students have adequate and equitable opportunities for learning.”

This critique is based on arguments developed for current litigation claiming that large numbers of children across the state attend schools without adequately “trained teachers, necessary educational supplies, classrooms, even seats in classrooms, and facilities that meet basic health and safety standards” and they are therefore “deprived of essential educational opportunities to learn.”¹⁸ The complaint seeks to hold state officials accountable for the conditions under which California students are expected to learn, particularly

¹⁷ Jeannie Oakes, Gary Blasi, John Rogers, co-authors of “Accountability,” 2003 in the forthcoming Kenneth Sirotnik, ed., *Accountability Run Amok*, New York: Teachers College Press

¹⁸ *Williams v. State of California, First Amended Complaint for Injunctive and Declaratory Relief*, p. 6

for the “shocking conditions in schools most likely to be attended by poor children and children of color.”¹⁹ The argument is that conditions, resources and opportunities matter and make the difference in how well students perform. Systems that hold students, teachers and schools accountable for below level proficiency assume that all that is needed to motivate teachers and students to improve learning are rewards or punishments based on test scores. It is further argued that it is the state that is responsible for ensuring that districts have the facilities, qualified teachers, materials and resources that would enable schools to meet the state standards.

An empirical study of teachers, principals and district administrators reinforces some of the argument that the current system of accountability does not adequately support their efforts to address inequities in student achievement.²⁰ Teachers in the study indicated feeling “frustrated and overwhelmed” at the same time they recognized the need for accountability. They appreciated certain aspects of California’s system, such as the standards component, while complaining about others such as the testing component whose results often arrive too late in the school year to be of much use to teachers. They also regarded the rewards and sanctions elements as arbitrary and biased. Testing and test preparation, it was found, often displaces other instructional activities and leads to a narrowing of the curriculum and diminished satisfaction with their work.

The study underlines the concern that current accountability policies, while drawing attention to discrepancies in achievement across groups of students and holding schools responsible, do not include effective solutions. Such solutions lie beyond simply testing students; they require a deeper investigation as to why inequities persist. Moreover, many educators complained they lacked proper tools to address issues of equity, including time, resources and administrative support. Holding all schools to the same expectations while failing to address disparities in resources, materials and funding was viewed as inappropriate. Finally, insofar as current accountability policy assumes schools as the primary locus of reform, this study shows that districts, through their policies and programs, have the power to facilitate or hinder school improvement. The study recommends that the state recognize district influence in the implementation of policy, and that additional diagnostic assessments, not just tests, be added to measure improvements, so as to provide immediate feedback, helping teachers to focus instruction and facilitate student learning.

This study and others, summarized in a Policy Brief by the Consortium for Policy Research in Education, published in February 2004, suggests that key concerns facing California’s assessment and accountability systems are the following:

- The need for continued focus by teachers and administrators on improving student achievement, particularly in low-performing schools
- Identification of schools as low-performing and publication of API scores and ranks, while useful, should not be associated with the use of rewards or sanctions tied to test scores.
- Low-performing schools currently lack the capacity to improve student learning in part because teachers lack the ability to use data effectively, despite additional state assistance.

¹⁹ Oakes etc., op.cit, p.1

²⁰ “Educators Respond to Accountability: Voices from the Field,” PACE, University of California, Berkeley, 2004

Key Question Remains

A key question remains: How useful or effective are accountability and assessment systems in improving student learning. There are clearly benefits of testing, and standardized tests can do a good job of indicating whether students have mastered basic skills. Tests have limits, however. They are of little use in assessing creativity, insight, reasoning and the application of skills to new situations. Such matters can be assessed, but not easily and not in a standardized fashion. “To judge schools exclusively by their test results is, therefore, to miss much of what matters in education.”²¹ To truly assess whether schools are improving student learning, Rothstein argues that schools need to include the judgments of experts who visit schools, review student work, evaluate the quality of curriculum and teaching, rate school climate, draw conclusions about the effectiveness of school leadership and make determinations about whether school resources are being devoted in a balanced fashion to all the goals we have for learning—both those that are easily testable and those that are not.

²¹ Richard Rothstein, “Testing Our Patience”, *The American Prospect*, February 2004, p. 45

A Case Study: The Fallout of the Accountability System

Both the California and the federal No Child Left Behind accountability policies operate under the assumption that disaggregated achievement data is a mechanism to highlight inequities and close gaps. Moreover, state and federal sanctions are tied directly to a school's ability to meet performance targets by subgroup, and not simply schoolwide. While such a focus on subgroups has increased educators' awareness, it occasionally results in a case of misguided attention.

Willow Elementary School presents such a case. While still below the state target of 800, Willow has made significant gains in overall student achievement, raising its API score from 557 in 2000 to 664 in 2003. The school's Latino and low SES subgroups have made considerable progress as well, moving from 459 to 599, and 493 to 628, respectively. Willow's other two subgroups, white and Asian students, have consistently outperformed their peers, with scores of 699 and 780 respectively, in 2003.

In 2002, Willow's Asian subgroup scored 741, yet missed their growth target by three points. A similar situation occurred the year before when the white students, Willow's second highest subgroup, improved but failed to meet their target by one point. NCLB legislation includes a sanction against any school that fails to meet a subgroup growth target two years in a row. Thus Willow was told it would have to send out a letter informing parents that the school was underperforming and that they could choose to transfer their children to a different school, at Willow's expense.

In many ways, Willow represents a school that has embraced the system of accountability, adopting state and district standards and implementing a successful reading intervention program. The school had recently been portrayed as a success story in the local media and held up as a model for other struggling schools in the district. Understandably, the news of this letter came as a shock to administrators and teachers.

Educators felt demoralized by a misdirected "low-performing" label, and frustrated by continued pressures from district, state, and federal policymakers despite improved test scores. As the school's principal remarked, "It just makes you feel like all your effort is for nothing." The principal was particularly concerned about how this would affect his staff:

Well, I think it's going to most affect the teachers because we work so hard. Just like, all the pressure we experience, we've totally come together as a team and we're feeling really good about ourselves, and then we get this news that we didn't do good enough. And all the other subgroups just did amazingly well, except the highest one didn't go up enough, you know? And because of that, we're being labeled a low-performing school. I can't even believe the idiocy of that.

The educators we interviewed were acutely aware of the fallacy of the policy. The superintendent described the situation as "absolutely ridiculous," calling this "the fallout of the accountability system." While attention to student performance by subgroups is a necessary step to addressing inequities in achievement, it is important to recognize that a singular focus may not allow for the complexities of school reform. The intentions of accountability reform had backfired; now the policy was drawing attention to the highest-performing students. And in its haste to hold schools accountable for raising all students' achievement, the policy left no room to recognize one school's efforts to close the gaps.

In a further twist of irony, Willow was later informed that it would not actually have to send out the letter, following the state's recalculation of API scores in 2003. With an adjustment of a few points, the school had met its subgroup targets and was now deemed a success.

Source: *Educators Respond to Accountability: Voices from the Field*, PACE, University of California, Berkeley, 2004, page 47.