

Chapter 6

Assessment

Highlights of Findings

The SAT-9 has made a significant impact on schools and teachers.

Teachers are highly aware of the SAT-9 and its importance. At many schools, the influence of the SAT-9 goes beyond test preparation and extends into the realm of shaping the curriculum itself.

As much as it may drive instruction, the SAT-9 has been the cause of much anxiety at the school level.

Principals and teachers expressed grave concerns about overreliance on the SAT-9. A lack of alignment between the SAT-9 and the curriculum is one major area of concern; a lack of alignment between the SAT-9 and content standards is another.

Many teachers feel that they are being compelled to “teach to the test” and that this may harm students.

Some teachers believe that ultimately, teaching to the SAT-9 will negatively affect students’ understanding of and appreciation for mathematics, as the test focuses on breadth rather than depth and does not sufficiently measure different types of mathematical achievement, such as conceptual thinking.

The augmented section of the STAR program caused particular anguish among teachers and students in spring 1999.

Although the augmented portion of the STAR program may have been more aligned with the state standards than the base SAT-9, many teachers felt that the augmented items were grade-level inappropriate and unfair to students, given the preparation they had had. Some teachers, however, indicated that they planned to adjust their curriculum coverage so as to better prepare students for the augmented items.

The quantity and timing of assessments can be problematic.

Several teachers and principals commented that too many assessments were taking time away from instruction. Also, the time at which any given assessment is administered plays an important role in how much of the content students have covered. Some teachers remarked that the SAT-9 included items that were not taught until mid- or late spring, after the test was administered.

Background

As with content standards, assessment in California over the past decade has had a rocky history. In the early 1990s, California implemented its first performance-based assessment system, the California Learning Assessment System (CLAS), specifically designed to measure students' mastery of curriculum laid out in the state *Frameworks*. However, in 1994, after just one year, funding for the test was vetoed by the governor for a combination of political, technical, and ideological reasons. In 1995, the state enacted the California Assessment of Academic Achievement Act (AB 265), which provided districts with funding to administer tests selected from a state-approved list.

Then in 1997, the Standardized Testing and Reporting (STAR) program was enacted. STAR, which was motivated by a perceived need for a statewide, comparable measure of academic performance for districts and schools that could report individual scores for all students, required all districts to administer the same nationally normed, “off-the-shelf,” basic-skills, standardized test. The test selected as the centerpiece of the STAR program was the SAT-9 (Stanford Achievement Test, Ninth Edition, Form T), published by Harcourt Brace Educational Measurement. The STAR program, still in force today, required virtually all students in grades 2–11, including English language learners, to take the SAT-9 each spring.

Meanwhile, as part of the statewide Standards-Based Accountability System, most districts were required in 1997–1998 to implement multiple measures of assessment for at least one grade level in each of three specified grade spans. The SAT-9 had to be one of the measures (as specified by the STAR program), but districts were relatively free to choose the other measures, provided that certain criteria were met and that the different measures were combined (to determine student proficiency) in accordance with state guidelines. For mathematics, many districts elected to develop or purchase criterion-referenced or performance-based assessments to meet the multiple measures requirement (Guth et al., 1999).

In 1999, however, the Public Schools Accountability Act (PSAA)—the enactment of SBX1—replaced the Standards-Based Accountability System and its multiple measures requirement. Under the provisions of the PSAA, the SAT-9 is currently the sole indicator being used in a statewide index designed to rank schools' performance and determine their eligibility for a rewards and intervention program. Until other indicators of academic performance are deemed valid and reliable, the SAT-9 will remain the sole measure of student achievement.¹ As such, it has become a truly “high stakes” test.

¹ In spring 1999, the test was “augmented” with extra items designed to assess student mastery of the content standards adopted by the State Board of Education in 1997. Student achievement on these items is measured separately from the base test. A study conducted by William H. Schmidt of the Third International Mathematics and Science Study (TIMSS) Center found that the base (mathematics) SAT-9 is not aligned with the California mathematics standards (Boser, 1999).

The Impact of the Stanford-9

The SAT-9 has made a significant impact on schools and teachers and in some places appears to drive curriculum and instruction.

Although, in theory, content standards (discussed in the previous chapter) should play the most important role in shaping curriculum and instruction, data suggest that assessment—and the SAT-9 in particular—actually carries more force. On the survey, 71% of fourth-grade teachers agreed strongly with the statement, “There is a school-wide effort to improve student mathematics achievement on the SAT-9.” In contrast, only 51% of fourth-grade teachers agreed strongly with the statement, “There is a school-wide effort to implement our district mathematics standards.” At the eighth-grade level, the figures for the two items were a bit closer together, but the SAT-9 still “won” over standards, with 80% agreeing strongly about the SAT-9 but only 70% agreeing strongly about the district standards.

Interviews with school-level personnel confirmed the importance of the SAT-9. Numerous principals and teachers spoke about “living and dying by the test scores,” focusing professional development efforts on improving test scores, pacing instruction so that teachers can “strategically prepare” the students for standardized tests, and “anxiously awaiting” the SAT-9 results. (Interviews were conducted before the scores were released.) One principal explained that “the SAT-9 has been the catalyst” for changes occurring in her school; “Other state policies,” she continued, “have had nowhere near the same level of influence.”

Indeed, in answer to the question, “Did you do anything special to help your students prepare for this year’s SAT-9?,” the vast majority of teachers interviewed answered in the affirmative. A few of the teachers focused on basic skills or on particular content areas as part of this preparation. One eighth-grade teacher, for example, explained that her school had identified fractions and decimals as an area needing improvement on the test, “so we did a lot of review on that concept.” Two fourth-grade teachers (both at the same school as one another) mentioned involving parents by speaking with them and telling them “we needed to help students prepare” or by sending letters home telling parents what skills were being tested.

More common responses, however, included work on “test-taking skills” (for example, in taking multiple-choice tests) and the administration of practice tests. As one teacher put it, “My main focus was teaching them how to take a test, as opposed to how to take this *particular* test.” Another teacher, similarly, explained, “My focus was not on math as much as on how to read the questions.” Several teachers mentioned the use of test-preparation booklets/materials, although in more than one instance, these materials had not arrived in time to be used for the current school year.

The amount of time spent specifically on SAT-9 preparation was variable. Roughly one-third of the teachers who were interviewed said they'd spent two to three weeks; about another third said one to two months or one day per week all year long. A few teachers reported that they had worked on SAT-9 preparation all year.

At many of the schools visited, the SAT-9 had an impact well beyond preparing students to take the test, extending into the realm of shaping the curriculum itself. Without prompting, many teachers mentioned the SAT-9 in their answers to questions about their “general approach” to teaching mathematics or about documents and policies that they felt had had an impact on their teaching. “The thing that jumps to mind is the STAR-9 testing,” replied one teacher; “the greatest impact comes from the Stanford-9 and [another assessment used in the district],” stated another. Responses such as these, along with “preparing students to take standardized tests,” were fairly typical.

Moreover, several principals stated unequivocally that the SAT-9 will “drive the way we teach” or had already done so. (Some interviewees acknowledged that assessment in general, rather than the SAT-9 alone, is the driving force.) The following comments were made by principals in three different districts:

We did a curriculum map last year related to the SAT-9. As a result our program has been skills based.

SAT-9 played a large part [in influencing mathematics instruction at the school]—fortunately and unfortunately. You want to teach the students what they will be tested on.

We use the make-up of the SAT-9 to determine what parts of the curriculum we should stress. For example, if there are more estimation problems on it we will cover that more next year.

Some teachers, as well, made comments about the influence of the SAT-9 over their curriculum or their instructional practices. “The test influences what I teach,” explained one teacher; “I try to cover all the areas that will be on the test,” she continued. In a different district, a teacher remarked that after the students had taken this year’s SAT-9, she asked them what they did not know on the test; they indicated geometry, so next year she intends to bring that in earlier. More generally, this same teacher stated, “If the SAT-9 is a test of skills, not theory, then we might as well continue to teach that way.”

Perceived Problems with the SAT-9

As much as it may drive instruction, the SAT-9 has been a source of much anxiety at the school level. Principals and teachers expressed grave concerns about overreliance on the SAT-9.

In response to the open-ended survey question about policies that have hindered mathematics teaching, 14.1% of fourth-grade teachers and 11.6% of eighth-grade teachers discussed assessment. Responses relating to assessment formed the second-largest category of responses to the question at the fourth-grade level, and the third-largest category at the eighth-grade level.

Many teachers commented simply that they felt there was too much emphasis on the SAT-9, on standardized testing, or on test results. Some teachers did indicate a belief that assessment as a measure of accountability is important— they just think that the SAT-9 may not be the most appropriate measure, particularly if it is the *only* measure. In response to the survey question about helpful policies, one eighth-grade teacher wrote, “Our district and school has focused on student learning and assessment has become a key issue. We look at assessment from many perspectives, not just testing.” And an eighth-grade teacher who was interviewed commented:

I would hope we're being held accountable. The problem I see is that I don't think it's [the STAR test] the one way you test for that. I think it should be just one of a variety of things. But I definitely think we should be held accountable for student performance. If not, we're not doing our jobs....I just don't think it [accountability] should be measured with one set of tests, and that's it. The kids I have...are good kids; they came in with good scores, they'll go out with decent scores; they probably could have done that no matter whether I did a good job or not. On the other hand, you can get kids that are ill-prepared, and you know, how much you can help them improve — I don't know that anybody knows, is that 5 percentage points? Is that 25 percentage points? I guess we're all wondering, what's going to be the measure of achievement? So, that's all a little iffy when the test is the thing.

The primary concerns that teachers expressed about the SAT-9 and its effects on instruction—and on students—are discussed in the following sections.

A lack of alignment between the SAT-9 and the curriculum is a major area of concern.

One frequently cited concern about the SAT-9, as discussed in the chapter on curriculum materials, was that curriculum materials are not aligned with the test. “I’m seeing that my students struggle with standardized testing because the curriculum adopted program does

not completely coincide. They have difficulty with transferring information learned while taking state test,” wrote one fourth-grade teacher on the survey.

Many teachers who were interviewed expressed a similar sentiment. “The Stanford-9 test material is not in our curriculum!” bemoaned one eighth-grade teacher. Another spoke of how the SAT-9 was a “more traditional” approach that does not mesh with the curriculum. A fourth-grade teacher had even more to say about this:

The new adoption for the district — there’s an obvious philosophy behind it that it should be hands-on...My biggest complaint with the hands-on is that [students are] not tested that way. It’s like they [the district] want us to use hands-on materials, but then they test us in a much more traditional way, and the students, at least in this school, have a very hard time making that connection, you know, applying the hands-on stuff to the test. [And the test] is what the district’s looking at...Regarding the district and the state, teachers are getting mixed messages about hands-on versus seatwork. I don’t get a consistent message. No one fully explains to you how you’re supposed to prepare kids for tests.

One principal who was interviewed said that there had been much anxiety in her school over the STAR program; she said that the teachers were worried that the kids were being tested on topics not taught. A principal in a different district made a similar comment, about teachers seeing “a discrepancy” between things on the test and things that are taught. Several interviewed teachers confirmed this. “The test doesn’t assess what’s going on here,” stated one teacher; “The SAT-9 is not a good judge,” said another.

Many teachers feel that they are being compelled to “teach to the test,” a particular problem if the test lacks balance and is not aligned with the standards.

As suggested by the remarks from those who say that the test is driving curriculum, it appears that many schools and teachers are adapting instruction to fit the test. But many teachers strongly object to the idea of “teaching to the test,” and believe that the overall effect on students will be negative. “Teaching for ‘the test’ drives the curriculum, in some areas to the detriment of what the students need,” wrote one teacher on the survey.

Again, teachers who were interviewed echoed this sentiment. As one eighth-grade teacher stated emphatically, “The SAT-9 is going to have a negative impact. It really controls teaching and what is taught.” Another interviewed teacher said that although he does not “believe in teaching to a standardized test,” he feels “tugged in that direction, because everybody thinks it’s important,” and thus has to “honor it.”

Some principals also expressed concerns about curriculum driven by assessment. One principal commented that looking at test scores might help improve the scores, but that this

did not necessarily mean improving the curriculum. Another principal said that she worries that as teachers teach more and more narrowly to the tests, important things are getting left out of children’s education. Previously, this principal remarked, she would have felt accountable to parents to give children a well-rounded education, but she now feels accountable to the district (who, in turn, is accountable to the state) to provide high scores. She thought that this sometimes gets in the way of giving students the best possible education.

Teachers helped provide an answer to the question of what, exactly, might be getting left out of children’s education as a result of the emphasis on the SAT-9. As with instructional practices and curriculum materials, some teachers expressed the concern that the test lacks balance between computational mastery and conceptual understanding and between depth and breadth, and thus that “teaching to the test” inhibits a well-rounded mathematics instructional approach. Survey comments along these lines—each from a teacher in a different district—included:

[from a fourth-grade teacher, cited as obstacle] Trying to teach conceptually when we are responsible for the students doing well on a standardized test that is traditional.

[from an eighth-grade teacher] The concern should be depth and understanding. Assessment tools need to address other intelligences. CLAS had the right idea. We need a TRUE multiple measure, not another multiple choice test.

[from a fourth-grade teacher, cited as obstacle] Pressure to “teach to the test” and not have students explore and enjoy mathematics as much as I would like them to.

[from a fourth-grade teacher] The time spent skimming over topics to prepare students for standardized tests could have been better spent by focusing on interesting concepts more thoroughly.

[from an eighth-grade teacher, cited as hindering policy] The emphasis on the SAT-9! I am encouraged to spend time on too many topics so students don’t get enough depth to remember topics so what they know this week they forget.

Another major concern that many people voiced about the SAT-9 is that it is not aligned with content standards. For example, one principal said that “we have no measure” for determining if a student meets the district standards, implying that the SAT-9 does not serve this purpose. Another principal mentioned that there had been “some resistance to the SAT-9 because it is not aligned with the standards.” She expounded further:

Do [the district math] standards align with the standardized tests that [students] have to take? No. They don’t. And that’s very frustrating for math teachers. What we’re teaching and when we’re teaching it, and when they take the standardized tests and they see that something is on

there that they haven't taught yet — it's very frustrating. ... Aligning our state testing with our standards is really important, so we don't have that frustration.

Indeed, some teachers particularly objected to the idea of “teaching to a test” that is not aligned with the standards. Two interviewed eighth-grade teachers commented:

I get the impression from the 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. Because now they [the state government] are offering extra money tied to teachers whose test scores are high. And, so that speaks very loudly that...it doesn't matter about the content, let's teach to the test....I'm not going to, but that's what I'm hearing, and I'll bet you that, in time, the department will force me to do that....I think the standardized test that we have to take gets in the way. Because it forces me to teach to the test, instead of teaching to what the standards are.

There's a lot of pressure to make sure students perform well on [the SAT-9]. And personally, I think if the curriculum is strong and you teach the curriculum, then you don't have to worry about the individual test. But, I'm kind of shouting out in a field by myself on that. Or, at least, there are a lot of teachers shouting out there, and other people aren't listening. And I just fear that we're moving too much toward teaching to a test. It's not ever been stated that way, but I think it's moving in that direction. I avoid it [teaching to the test], thinking that the strength of the curriculum will do the job. And, I don't know what I'll have to do if the results aren't good, and I have to revise what I do. Because, I think, then the task is, change the curriculum...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. And I don't think we're anywhere near there yet....

Of course, there is the further issue of which standards the test should be aligned to, given that (as discussed in the chapter on standards) different sets of standards—district, state, national—may not be aligned with each other. One fourth-grade survey respondent remarked, “There is a discrepancy between the need to cover all possible test topics to improve test scores, and the NCTM standards that emphasize thoroughness and deeper understanding of concepts and number sense.”

Although the augmented portion of the STAR program may be more aligned with the state standards than the base SAT-9, the use of the augmented test in spring 1999 caused considerable anguish among teachers and students.

At least in theory, the use of the new “augmented,” standards-based sections of the STAR program may alleviate some of the concerns that people have about lack of alignment

between the test and the standards. However, it appears that considerable progress remains to be made with the use of these new sections. According to sources within the California Department of Education, some of the augmented items (as administered in spring 1999) failed to meet technical standards of validity and reliability.

Moreover, several people objected to the augmented sections of the 1999 test on the grounds that they were unfairly difficult, especially given the level of preparation most students had had prior to the test. On the survey, one fourth-grade teacher wrote that augmented test was “despicable.” “After hours of dreary testing,” she continued, “students are made to feel ignorant of things they have never laid eyes on. I am disgusted.” Another fourth-grade teacher reported that teachers had not been informed about “the new augmented portion of the math test that was added” until shortly before the test was administered and that there had been “no helpful information to aid or guide us.”

Other survey comments suggested the test’s content was grade-level inappropriate. Many of these comments did not mention the augmented sections *per se*, but, given other remarks that were made, it seems likely that the augmented sections were the basis for the comments. The following remarks were made by fourth-grade teachers in three different districts:

[cited as hindering policy] Rewriting requirements to meet STAR (which are not reasonable to begin with), which essentially want me to push 4th graders into 6th grade math without experiencing 5th.

[cited as hindering policy] State tests should test concepts taught at this grade level.

The “Star” testing is inappropriate for the “average child”—Great info for the students that excel in given areas/topics. I question the validity of results.

Similar findings came from interviews. One principal remarked, “the augmentation portion was a bust”; she said that the test “set the students and teachers up because the expectations were not matched by what students found on test.” And the following remarks were made by interviewed teachers in two different districts, the first one an elementary-school teacher and the second one a middle-school teacher:

I was really upset by the augmentation test. The students were asked to work with negative integers. I didn’t teach them that.

The SAT-9 tests a lot of stuff that they haven’t even learned...The problem is that we’re supposed to be aligned with the state test. And so, that means basically we need to advance all our students before they’re ready....The seventh graders had to take this test, the STAR test...While they were taking it, I could just see the frustration on their faces, and I was like, what’s going on? ... [I

realized], oh my gosh, they're so frustrated because this is the stuff I'm teaching my eighth graders right now, but my seventh graders haven't even seen this material yet.

This teacher said that as a result of this experience, next year he plans to move content down from the eighth grade to the seventh grade to the “best of his ability.” Similarly, an interviewed fourth-grade teacher said, “I don’t believe in teaching to the test but it’s not fair for a child not to have exposure to what’s on the test.” She indicated that next year, she will add new topics to her curriculum—those on the augmentation test—so that students have exposure to them.

In this way, then, the use of the augmented portions of the STAR program may indeed be having the effect desired by the state: they seem to be spurring at least some teachers to teach particular content at levels they otherwise would not have. To the extent that this content is indeed aligned with the standards, then the test is encouraging standards-based instruction. As one principal put it, “[The augmented test] has really been an issue with our math teachers, because they feel that it’s out of reach of most students. But maybe that’s the purpose of it: make it within reach.” However, this same principal also stated that there had been “a lot of resistance” among teachers to changing their curriculum to match what was on the state augmented test.

Quantity and Timing of Assessments

Testing takes time away from instruction.

Another area of concern with regard to assessment is the amount of class time needed to administer and prepare for tests. Many teachers felt that this time could be spent in more instructionally valuable ways. On this matter, the SAT-9 was viewed as only one of the culprits; other assessments, such as those required at the district-level, also were partly responsible. Survey comments along these lines included:

[from a fourth-grade teacher] There have been a large number of tests required this year that took away from teaching time and covered areas not presented in our current text. There should be a more relevant, valuable, and enjoyable way to assess and educate students.

[from an eighth-grade teacher] Too many standardized tests given in fourth quarter cause loss of teaching time and promote apathy in the students.

[from an eighth-grade teacher] If you look at the amount of time taken by state and district assessments you lose about 5–10 days of instruction.

A principal who was interviewed also expressed the concern that too much time was being devoted to testing, and that it was cutting into instruction time. And an interviewed teacher in a different district said he thought that the district assessment, given three times over the course of the year, “was a little much”:

We lost three instructional days, plus whatever preparation we were doing for it. And then also, it took some time to grade the papers, all that kind of stuff, which took away from my preparation time as well....So I thought it was a little much...to do three of them; I felt it would be better if it was just one.

Some interviewed teachers said that they had stopped what they were doing in order to prepare students for assessments (including the SAT-9), and a few of them resented having to do this. As one teacher put it, “It [test preparation] slowed me down with respect to my regular instruction.”

The time at which an assessment is administered also plays an important role in how much of the content students have covered.

Some teachers voiced concerns not only about the amount of time required to prepare for and to give assessments, but the particular scheduling of these assessments, as indicated by the following interview comments from two teachers at one school:

I mean, it's really hard, because, like, we'll get a test coming up, a [district] performance-based assessment test, coming up, and I'll look at it, and I'll go, “Oh, gee, we haven't even covered this yet.” So I'll have to stop what I'm doing, cover this material, so that they can do well on the performance-based assessment test. And then go back to my regular material.

The district has had...performance-based assessments that we had three times this year...And I have no trouble doing performance-based assessments, but when it comes from the district, it doesn't necessarily fit with what you're doing at the time. I'd rather have an assessment that goes along with what they [students] are doing...It was like, just take this chunk out of time, and do this thing that's not associated with what you're teaching.

Another scheduling concern is that some assessments—the SAT-9 in particular—are administered before students have been exposed to all the content in the assessments. One interviewed eighth-grade teacher stated that although the SAT-9 was given in the early spring, it focused on the last third of the year's curriculum, and the class simply “hadn't gotten to a lot of those topics yet.” Another eighth-grade teacher, interviewed toward the end of the school year, said that her class had covered several more standards since the test was given, as a result of the way the book was set up. She hypothesized that if her students could “take the test today, they could get at least ten more right.” As it was, however, she

stated, “The SAT-9 was extremely frustrating — it was a whole week of upset and tears” for her students, whom she said are among the best at her school. She teaches five gifted classes.

In the Next Chapter

As shown in this chapter, the SAT-9 has made a significant impact on instruction, as teachers are eager to help their students do well on this high-stakes test. However, teachers’ good intentions alone may not be sufficient to raise student achievement. Even if student achievement on the SAT-9 does improve, achievement on measures of assessment that measure different types of mathematical skills and abilities might not. The implementation of meaningful instructional change that truly raises students’ understanding of mathematics might require changes in teacher preparation and professional development. These will be discussed in the following chapter.