



The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

Wayne E. Wright
Arizona State University

Drawing on interviews with teachers, this study examines the effects of a high-stakes standardized test (the SAT-9), on a large inner-city elementary school in southern California with a high English Language Learner (ELL) student population. Specifically, the study addresses the following questions: (a) How much of an emphasis is there on the SAT-9? (b) Do teachers believe the SAT-9 is a fair and valid measure of teaching and student learning? (c) How has the SAT-9 affected the curriculum taught to students, especially ELL students? (d) How has the SAT-9 affected the teachers and students? And, (e) has the SAT-9 improved teaching and learning at the school under study? The findings reveal that standardized testing has not resulted in higher quality teaching and learning in this school; rather, it has resulted in a narrowed curriculum and harmful effects on both teachers and students.

The majority of states are now using some form of high stakes tests to hold students, teachers, and schools accountable in an effort to improve teaching and learning. This study analyzes the effects of California's high stakes testing policy on an inner-city elementary school with a high English language learner (ELL) population.

Under the State's Standardized Testing And Reporting (STAR) Program (see <http://star.cde.ca.gov>), all students in grades 2-11 in California must take the Stanford Achievement Test, 9th Edition (SAT-9). The SAT-9 is a nationally-normed multiple-choice test produced by Harcourt Educational Measurement. Elementary students are required to take several separate language arts and math sections. Combined, the tests take considerable time. For example, the mandatory sections of the 2nd grade test take nearly 7 hours to complete.

California has created an Academic Profile Index, which uses results from the SAT-9 exam to rank schools throughout the state (Pyle, 1999).

Schools with low test scores are targeted and given growth rates to which they are held accountable. Low scoring schools that fail to meet the growth rate are scrutinized and are denied the monetary rewards given to schools and teachers who meet their growth targets. Newspapers publish SAT-9 scores and API rankings, subjecting schools and school districts to public scrutiny as well. Each consequence clearly contributes to the classification of the SAT-9 as a high-stakes test in California.

The purpose of this study is to analyze the effects of the SAT-9 on teachers, students, and the curriculum in an inner-city elementary school. This school serves a high population of English language learners and is situated in a low-socioeconomic neighborhood. Inner-city ELL students traditionally do not do well on standardized exams, yet are the students that testing proponents target under assumptions that high stakes tests will improve educational quality and academic achievement in their schools (Amrein, Berliner & Biddle, 2002;

Haney, 2000; McNeil, 2000; Ohanian, 1999 and Heubert & Hauser, 1999).

Specifically, the purpose of this study is to address the following questions: (a) How much of an emphasis is there on the SAT-9? (b) Do teachers believe the SAT-9 is a fair and valid measure of teaching and student learning? (c) How has the SAT-9 affected the curriculum taught to students, especially ELL students? (d) How has the SAT-9 affected the teachers and students? (e) Has the SAT-9 improved teaching and learning at the school?.

Literature Review

Today's testing movement has its roots in the eugenics movement. Early founders and advocates of IQ testing saw great utility in the ability of tests to sort out the so called "feeble-minded" children (and adults) for purposes of social control (Gould, 1996; Grissmer, Flanagan, Kawata, & Williamson, 2000; Heubert & Hauser, 1999; Sacks, 1999; Stoskopf, 2000). When IQ tests made their way into schools, educational practitioners found them to be useful tools for quantifying notions of merit and aptitude, sorting children, and disbursing educational resources accordingly. Tests became more popular as did society's reliance on tests to hold schools accountable.

Today, we test more students, with greater frequency, and with a larger number of tests than during any other time in the history of the United States. Over the last three decades we have also increasingly relied on tests, to which severe consequences have been attached, to reform our schools. These tests are known as high stakes tests (Sacks, 1999; Kohn, 2000). Advocates of testing argue that attaching stakes to tests is necessary to hold schools accountable, reward high performing schools, and identify failing schools so they may be targeted for extra help. This is a key element of President George W. Bush's education plan (No Child Left Behind ACT, 2001). However, researchers and numerous others have false notions that accountability measures such as high stakes tests actually yield increases in academic achievement.

There is mounting evidence that gains on state tests are not necessarily indicators of higher achievement. An experimental study by Koretz, Linn, Dunbar and Shepard (1991) revealed that performance on a high-stakes exam did not generalize to other tests for which students had not been specifically prepared. Klein, Hamilton, McCaffrey, and Stecher (2000), investigated the performance gains celebrated in Texas. They compared the Texas Assessment of Academic Skills (TAAS) scores with the scores taken from the National Assessment of Educational Progress (NAEP) and found the dramatic increases in TAAS

were not evident on the NAEP as had been previously purported. Additionally, while the TAAS illustrated that performance gaps between whites and students of color were narrowing, NAEP scores showed that the gaps were widening. Amrein and Berliner (2002) found a similar pattern as they examined increases on 18 states' high-stakes exams and patterns on other tests that tested similar knowledge constructs (e.g., SAT, ACT, NAEP, and AP exams). All researchers found that significant increases on high-stakes exams did not transfer over or generalize to these other exams, challenging the notion that high stakes tests caused increases in academic achievement. High-stakes accountability systems can and do get results (i.e., increased test scores), but the results are not "particularly deep or lasting" (Fullan, 2001, p. 220). Rather, the results are artificial.

Given that high stakes tests have not evidenced authentic learning gains, scholars express deep concerns over the negative and harmful results of testing on schools, teachers, students, and the curriculum (Haney 2000; Klein, Hamilton, McCaffrey, & Stecher, 2000; McNeil, 2000; Valencia & Bernal, 2000). Sacks (1998) contends, "Focusing exclusively on measurements and accountability may have [had] precisely the opposite of [their] intended outcomes" (p. 93). This was a major finding in the work of McNeil (2000). McNeil studied the effects of high stakes tests in classrooms in Texas—the state where dramatic increases in test scores were dubbed the "Texas Miracle." McNeil revealed that school reform efforts that which centered on testing, greatly distorted the educational experiences of students in urban schools. She found that as schools focused more and more on test preparation and teaching to the test, test scores increased, meanwhile the quality of teaching and learning was both compromised and depreciated.

In an in-depth examination of elementary schools in Arizona, Smith, Edelsky, Draper, Rottenberg, and Cherland (1989) found that external (i.e., state and district mandated) high stakes testing resulted in additional negative effects in the classroom. They found that:

1. Testing reduces the time available for ordinary instruction.
2. Testing affects what elementary schools teach. In high stakes environments, schools neglect material that the tests do not include.
3. Testing encourages use of instructional methods that resemble testing.
4. Testing affects school organizations by placing general boundaries on placements and instructional opportunities.

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

5. Testing has hidden structural effects on ordinary instruction.
6. By teachers' definitions, testing affects pupils.
7. Testing affects teachers. (pp. 267-275)

In addition, Haney's (2000) investigation of the "Texas Miracle" revealed that poor and minority students dropped out of school at significantly higher rates as pressures to raise test scores intensified. Thus, increases in the dropout rate led to higher aggregate test scores. Test proponents in Texas touted increases in scores as proof that high-stakes tests worked to improve academic achievement. What was more plausible of an explanation, however, was that the students who dropped out of school took their low-scores with them. A simple change in the sample of students who were tested caused the increases in composite test scores - actual increases in achievement did not. The students who remained in school did nothing but remain in school to improve their composite test performance. As Heubert (2001) maintained, "When low-achieving students are not part of the test-taking population, the pass rates of those who remain will be higher-even if the achievement of those who actually take the test has not improved" (p. 4). In fact, Haney (2000) and McNeil (2000) have even documented incidences in which school personnel in some Texas schools actually prevented low-achievers from taking the tests in order to inflate test scores.

Finally, the issue of the inclusion of language minority students who are not yet fluent in English, has been a major concern raised by many scholars, educators, and others in the testing debate. Language minority students frequently live in low socio-economic neighborhoods and attend inner-city or rural (migrant worker area) schools - neighborhoods and schools disproportionately impacted by high stakes tests. For example, "50% of the schools scoring in the lowest 20% on the [Academic Profile Index in California] were in rural areas ... while the other 50% were in the inner cities. ... All of these areas are impacted by poverty and language barriers" (Grisham, 2001, p. 13; see also Wells, 2001).

The language barriers these students face inevitably affect the ways in which they perform on tests. English-language learners are usually expected to meet the same standards as native English speakers on high-stakes tests "before having made the improvements in instruction that will enable all students to meet the standards" (Heubert, 2001, p. 11). The language needs of these students are oftentimes overlooked in large-scale testing programs. More specifically, accommodating the

language needs of language minority students is an endeavor not often valued in states with English-only policies such as California. As such, "large-scale assessments are studded with problems that affect not only how equitably the achievement of LEP [Limited English Proficient] students can be measured, but how effectively LEP students' mastery of content [can be] assessed" (Kopriva, 2000, p. 5).

According to the U.S. Department of Education's Office of Civil Rights (Office of Civil Rights, 2000) high stakes tests should not have a disparate impact on ELL students. Education agencies must provide an education program that enables ELL students to master the knowledge and skills necessary to pass high stakes tests. The American Education Research Association (2000) position statement concerning high-stakes testing in PreK-12 education extends this point:

If a student lacks mastery of the language in which a test is given, then that test becomes, in part, a test of language proficiency. Unless a primary purpose of a test is to evaluate language proficiency, it should not be used with students who cannot understand the instructions or the language of the test itself. If English language learners are tested in English, their performance should be interpreted in the light of their language proficiency. Special accommodations for English language learners may be necessary to obtain valid scores. (p. 13)

Despite professional guidelines (and research findings that show the negative impact of high stakes tests), SAT-9 scores and API rankings are being used inappropriately in California. The use of the SAT-9 as the sole indicator of school performance and student achievement is leading to the neglect of the academic needs of ELL students (Garcia, 2001). Likewise, the API in California is misleading for schools with large numbers of English learners, particularly as educational programs for language minority students are being evaluated post-Proposition 227¹. Several scholars have negated claims that English-only programs are superior to bilingual education programs, for example. Using SAT-9 scores such claims simply cannot be substantiated (Butler, Orr, Gutierrez, & Hakuta, 2000; Mora, 2001; Thompson, DiCerbo, Mahoney, & MacSwan, 2002).

From the literature reviewed herein, it is clear that there are serious concerns among scholars and practitioners about the effects of high-stakes testing. Other than McNeil (2000) and Smith et al. (1989), however, few researchers have actually gone

into the schools. Few have been able to document first hand the negative effects of high stakes tests in classrooms. While this study is not as comprehensive as the work conducted by McNeil and Smith and colleagues, it fills a void in the literature. It takes a closer look into an inner-city elementary school in California with a high English language learner population and investigates the ways in which school personnel were forced to come to terms with California's SAT-9, its Academic Profile Index, and the burdens of teaching in a state with a high-stakes test.

Methodology

Those who understand the effects of testing are those who are closest to the tests — the teachers. Ironically, teachers' voices are rarely heard in the testing debate. Their views are often dismissed by testing advocates who argue that teachers oppose high stakes tests simply because they do not want to be held accountable; teachers are biased so their concerns about high stakes tests should not be warranted. Yanow (2000), on the other hand, argues, "To understand the consequences of a policy for the broad range of people it will affect requires 'local knowledge'- the very mundane, expert understandings of and practical reasoning about local conditions derived from lived experience" (pp. 4-5). Clearly, teachers possess this "local knowledge" and have the "lived experience" necessary to understand the effects of high-stakes testing. Qualitative research methodologies were utilized to collect the data needed to tap into this "local knowledge" (Bogdan & Biklen, 1992; Merriam, 1998).

This study was conducted at Alamitos Elementary School², where I taught for several years. My interest in selecting Alamitos Elementary as the research site was the result of an informal visit to the school two months prior to the formal interviews. During this visit, each conversation in which I engaged with teachers, pertained to the enormous pressures they were under to raise test scores. Wanting to get at the particulars and access the voices of teachers, I set out to generate a more comprehensive understanding of the topic, lending empirical evidence to the overall literature pertaining to standardized testing.

Description of the Research Site

Alamitos Elementary School is located in Southern California in a low socio-economic inner-city neighborhood. As is typical of other inner-city schools, Alamitos is overcrowded, has a large ELL population, and is considered to be an "underperforming school" due to low SAT-9 scores

and the fact that it did not meet all of its 2000 or 2001 API growth targets.

Alamitos Elementary School is in the Beach Cities School District (BCSD), one of the largest school districts in the state of California. Alamitos provides instruction year round to 1,178 students (but was originally designed for only 500) in grades K-5 on four different tracks (A, B, C, and D) and busses out an additional 1,400 students to schools throughout the District. The tracks are not based on ability per se, but make it possible to educate many students in the limited space available in the school. Students on each track attend school for 3 months, followed by a one-month vacation. By staggering the vacations (off-track time) of the four tracks, more classroom space is available for the students and teachers that are "on-track." At Alamitos, as in other year-round schools, this requires one or more teachers at each grade level to be a "rover," meaning that their class must rotate to a different classroom each month. While the tracks were not designed to be ability based, a series of gifted classrooms developed along Track C. The former Spanish bilingual program developed along Track D and part of Track A. A small Khmer (Cambodian) bilingual program developed along part of Track B. In effect, some racial and ability tracking resulted.

The student population is 63% Latino(a), 31% Asian (mostly the children of Cambodian refugees), 4% African American, and 3% other (including about five white students). Eighty-seven percent of the students are classified as English language learners; 67% are Spanish-speaking, 28% are Khmer (Cambodian) speaking, and the rest speak Lao or Vietnamese. Fewer than 9% of the students' parents completed college and fewer than 25% completed high school. All students in the school qualify for a free school lunch.

Alamitos teachers have an average of 6 years of experience, which is below the State average of 13 years. Fifty-five percent of the teachers are fully certified and the rest are in the process of completing their credentials.

On the SAT-9 in 2000, 78% of Alamitos students scored below the 50th percentile in reading, and 66% scored below the 50th percentile in math. This earned Alamitos a score of 2 out of 10 (10 being the highest) on the Academic Profile Index (API). When compared to schools with similar demographics, however, Alamitos scored 7 out of 10. The school did meet the overall API growth target but failed to meet the growth target for its Asian "subpopulation." As a result, the school and its teachers were denied monetary rewards, and Alamitos was deemed a "Performance Review School" by the state.

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

In 2001, SAT-9 scores improved overall with a 2% increase in students above the 50th percentile in reading, and a 3% increase in students above the 50th percentile in math. The school even succeeded in meeting its overall API growth target, and the growth target for its Asian subpopulation. However, this time the school did not meet its API growth target for its Latino(a) subpopulation, and once again the school and its teachers were denied monetary rewards.

Also disappointing was a drop in API rankings from 7 to 5 (out of 10) in comparison to schools with similar demographics. This drop is somewhat suspicious given test scores increased. Similar school API rankings are calculated based on parental data provided by the students. Of course, many children from immigrant families, especially younger children, are not fully aware of this information. Parent data were only reported for 54% of Alamitos students, for example. However, a school less than a half-mile away, which is essentially in the same neighborhood, was able to report parent data for 80% of its students, and was able to show lower percentages of education levels of its parents. This school earned an API ranking of 10 on the similar schools scale, even though its overall reading scores were lower than Alamitos' scores. Its math scores were higher, but the wide difference on the API similar schools ranking (5 as compared to 10) reveals that these rankings may depend more on how diligently the school collects and reports parent background information than how well students actually perform on the SAT-9.

Interviews

I conducted formal interviews with five experienced teachers at Alamitos Elementary School. I focused on teachers in the primary grades because that is where grade level testing begins. I decided to focus mainly on the 2nd grade teachers because that is the first grade level in which students are required by state policy to take the SAT-9. I included one 1st grade teacher because in addition to the state policy, in the district testing policy it is written that all 1st grade students should be tested as well. Lastly, I included a kindergarten teacher to determine if the SAT-9 was having an effect on students' first year of schooling.

I felt it was important to interview teachers who were (a) experienced, (b) knowledgeable, (c) viewed by their peers as dedicated and successful teachers, and (d) representative of the various instructional models in the school (i.e., former bilingual teachers³, English-only teachers, and gifted teachers). I made a list of several teachers who fit my criteria, narrowed it down to five, and then proceeded

to contact each participant. I explained to them the purpose of the study. When they agreed to an interview, an appointment was scheduled.

All interviews were conducted a few months before the SAT-9 was administered in 2001. I traveled to California to conduct the interviews in person. Interviews were guided by an interview agenda I developed prior to the interviews, but teachers were free to divert from the interview questions to topics or issues they felt warranted discussion. The interviews lasted 60 to 90 minutes, were tape recorded, transcribed, and coded⁴.

Interview Participants

Linda Veal is a kindergarten teacher, and has over 17 years of teaching experience (16 at Alamitos). Linda is certified to provide English language development (ELD) and specially designed academic instruction in English (SDAIE) instruction to ELL students. She teaches mostly ELL students in an English-only setting, although she does speak some Spanish and has a bilingual aide. Linda served as a mentor teacher for several years, and has provided training and assistance to teachers throughout the district in areas such as literacy, math, and classroom management. In addition to teaching, Linda serves at the district level as a Peer Assistance and Review (PAR) coach and is currently completing a Masters Degree in Early Childhood Education.

Hina Anjanar is a 1st grade teacher, and has been teaching for seven years (all at Alamitos). Hina is Eastern Indian, but lived for several years in Panama before her family came to the United States. Spanish is Hina's second language and English is her third. Hina is a bilingual teacher and taught in the bilingual program for four years before the district dismantled it after Proposition 227 passed. She currently teaches mostly ELL Spanish-speaking students in an English-only setting.

Bianca Gonzales is a 2nd grade teacher, and has been teaching for four years, all at Alamitos. Bianca is originally from Cuba and did not speak English when she began school in kindergarten in the United States. Like Hina, she is a bilingual teacher and taught in the bilingual program before Proposition 227. During the 1999-2000 school year, she taught in an experimental bilingual program established by the district for parents who signed waivers. This year she teaches ELL Spanish-speaking students in an English-only setting, but still provides some Spanish primary language support when her students need it.

Mary Lemon is a 2nd grade teacher and has taught for nine years (all at Alamitos). Mary is a passionate and hard working teacher, has a background in child development, and is deeply

aware of the physical, emotional, and educational needs of her students. She is certified to provide English language development (ELD) and specially designed academic instruction in English (SDAIE) to ELL students, and has a bilingual aide.

Nicole Soto teaches the 2nd grade gifted class called "Excel." She is originally from Canada, and has taught for over 20 years in different countries and at various grade levels. She has been teaching at Alamitos for over nine years, has a Masters degree in reading, and is certified to provide ELD and SDAIE instruction to ELL students. Nicole does not speak Spanish, but her husband's family was among the original inhabitants and landowners of California before it became a part of the United States. Despite the fact that Nicole teaches in the gifted program, the majority of her students are ELLs - the Excel program is designed to identify and serve gifted students in inner-city, low-socioeconomic schools.

Observations

My own experiences and observations in Alamitos and the district served as a secondary source of data. I experienced many of the issues described by the teachers at this site. In addition, I have been in the classrooms of the participating teachers many times; have had opportunities to see them teach and interact with their students; sat in staff meetings, grade level meetings, and curriculum planning meetings with them; and have had numerous informal conversations with them. My "insider knowledge" made it possible to verify the claims made by teachers regarding the changes in the school and their teaching as a result of the SAT-9.

Documents

A third source of data was generated from district and school documents. These included content standards and assessments; memos; handouts from meetings, trainings, and workshops; agenda and meeting minutes; working documents; announcements; and newsletters. Many of these documents were gathered while I taught in the district, and additional documents were obtained from the study participants and other teachers. These documents were carefully analyzed and used to triangulate or substantiate the statements made by teachers.

Findings and Discussion

Recurring themes emerged from the data and served to organize the discussion herein. Where appropriate, themes are illuminated in vignettes (see Erickson, 1986).

Emphasis on the SAT-9

All of the teachers commented about the heavy emphasis placed on the SAT-9. It is important to note that when the test was first administered in 1998, there was almost no emphasis on it whatsoever.

No test preparation materials were purchased, and no test preparation time was required. At a staff meeting, we were told not to worry about it and to tell our students to just do the best they could. Several teachers speculated that this was intentional. There was a desire to have low baseline data in order to show improvements in the subsequent years. Increases in year-to-year scores would substantiate claims that various school reform initiatives were effective.

In just three years, the emphasis on the test has become so great, that many staff members now sarcastically refer to the district as "SAT-9 Land." Nicole described the current emphasis on the SAT-9:

It's an avalanche. I think that is the best word to use. It's an avalanche. It's an avalanche that is sweeping common sense and good teaching practices, just bulldozing them out of the way. It has created such a test-driven, stress-driven situation.

Mary explains the high priority being given to testing and test preparation:

Everything that has to do with the test has been given such a high priority, that there is no priority any more but that...The bottom line question comes down to, "Well, what's going to help them do better on the test?" And if it's not going to help them do better on the test, well, we don't have time for that right now.

Linda said the emphasis on the SAT-9 is "incredibly more!" She added:

I would say it has quadrupled at least in terms of its importance, in terms of the amount of time we spend at staff meetings dealing with it, and with the consequences of it. It really is now high stakes.

Hina said, "And now that's all we're worried about. Test test test test test!"

And Bianca commented:

This year, I think all I care about is, "OK, am I using the vocabulary that the kids will see on the test?" All I think about [is], "OK, let's see what's on the test, and let me teach to the test."

The teachers' statements are supported by documents in which a large number of district and school memos and information packets addressing the SAT-9 were found. These documents include guides correlating district curriculum with test preparation materials and the SAT-9, staff meeting agendas, test preparation newsletters, and changes to district standards and assessments.

Bianca described meetings at the beginning of the 2000-2001 school year in which teachers were released from their classrooms to meet with

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

administrators to set professional goals. They were instructed to bring SAT-9 scores of students below the 20th percentile. In the past, teachers' goals might include such things as increasing their own knowledge on a particular instructional technique, differentiating instruction to meet the needs of all students, or improving the oral proficiency of ELL students. Now, teachers' goals were prescribed and driven by the needs to raise SAT-9 scores.

Mary mentioned that the school held assemblies geared toward teaching students the *dos* and *don'ts* of testing and motivating students to "test your best." Vignette 1 describes one such assembly:

"Test Your Best" Assembly

Twelve 2nd grade classes are crowded into the well-worn, dimly lit auditorium with only 200 seats. Two classes are seated on the floor in front of the stage. Teachers were doing their best to quiet down their students, many of whom were chatting loudly or bouncing up and down on the torn cushion seats. The commotion came quickly to an end as Mr. K., the young and hip school counselor leapt onto the stage, grabbed the microphone, and yelled in his Bob Barker voice, "Welcome to Test Your Best!!!!" A heavyset Latino boy sitting at a table on the side of the stage pushes play on a tape recorder and game show music comes blaring through the speakers. By now all students are sitting eagerly looking towards the stage. Mr. K. has gotten their full attention.

The music stops and Mr. K reminds the 2nd graders that in a few weeks they will be taking a very important test - the SAT-9. He tells the students that it is very important for them to do the best they can. He then tells the students that his 5th grade friends will be playing a game called "Test Your Best," to help teach them what things they can do to test their best!

"Contestant #1, come on down!" Mr. K yells. A Cambodian girl comes running down the left aisle of the auditorium. Mr. K turns to her and asks, "Sokha, what can you do to make sure you do well on the test?" Sokha answers in a shy, subdued voice, "Sit next to someone really smart and copy their answers?" Buzzzzzzzzzz! The sound boy sets off a loud buzzer with a big grin. The students laugh. Mr. K turns to the audience and using his counselor voice, asks "Why is her answer wrong?" About 30 little hands go up. A few

boys are jumping up and down going "Oooh! Oooh!" Mr. K picks a boy in the second row. "That's cheating!" the boy says emphatically. "That's right!" proclaims Mr. K. He then leads a discussion on why it is wrong to cheat, and how it is important to study and do it on your own.

Mr. K yells "Contestant #2, come on down!" A Latino boy comes running down the aisle. Mr. K asks, "Jonathan, what should you do the night before the test?" "Jonathan takes the mike and says, "Stay up all night and watch TV?" Buzzzzzzzzzzzz! Once again the kids laugh. Mr. K turns to the audience and asks, "What should you do the night before the test?" about 50 hands go up, waving eagerly in the air. Mr. K picks a girl seated in the left section of the auditorium. She mumbles the answer so softly no one can hear. Mr. K prompts her to speak louder, but once again she mumbles, and then lowers her head, suddenly not enjoying all the attention focused on her. A boy next to her yells out, "She said go to bed early!" "That's right!" pronounces Mr. K. He then switches to counselor mode, and tells the kids how important it is to get a good night sleep before the exam.

Mr. K then yells out "Contestant #3, come on down!" A Cambodian girl comes running to the stage. Mr. K asks, "Jennifer, If you are taking the test, and you don't know the answer to a question, what should you do?" Carefully reciting back her memorized line, she speaks slowly and softly into the microphone, "Mark a guess and go on to the next question." Ding ding ding ding ding ding! The giggling sound boy bangs a desk bell several times. The students break out in wild applause. "That's right!" exclaims Mr. K. He begins his counselor lecture on how important it is to mark a guess for every answer and to never ever leave any questions blank. He also stresses several times how important it is to only mark one answer for each question and to make sure to erase wrong answers really well, or else the computer will score it as "incorrect".

"Contestant #4, come on down!"... The show continues with a few more contestants. Near the end of the show the student helpers line up on the stage holding posters rested against their legs with the blank side facing the audience.

Mr. K yells out, "What's the first thing we learned?" Sokha holds her poster up over her head and Mr. K reads it to the students, "Don't cheat." Let's all say that! The students repeat loudly "Don't cheat!" "What's the 2nd thing we learned?" Mr. K. makes an exaggerated pointing motion to Jonathon, who holds up a poster reading, "Go to bed early." Mr. K reads it and has the students repeat it. This time, they repeat it a bit louder. "What's the 3rd thing we learned?" At Mr. K's signal, Jennifer holds up her sign, "Mark an answer for every question." Students enthusiastically repeat it back, some standing and yelling at the top of their lungs. This exercise is repeated, with signs and students yelling things like "Mark only one answer!" and "Eat a good breakfast!"

At the conclusion of the assembly, the students clap and cheer wildly. As they exit, several chant "Answer every question!" or "Only mark one answer!"

Prior to the assembly, motivational posters along the theme of "test your best" were created by students and posted all over the school, and remained up until SAT-9 testing was completed. Nicole mentioned some of the other "absurd" practices happening as a result of the emphasis on the SAT-9:

And then last year [there was] the added stress of "If all your kids show up every day for the test you get a field trip." Uh huh. It's like, "Oh good. I want the kid with chicken pox coming into my classroom to take a test? I don't think so!" You know, that was horrible. That was absolutely ludicrous; another sign of insanity. I had a chronically ill child with a horrible disease...how good was he going to do on the test? But hey, there was the carrot! "Get all your kids in to test, you get a field trip." For goodness sakes, a field trip! Shouldn't we be giving children in the inner-city particularly, field trips? They should have more than the children who are in the more affluent areas.

The implementation of Governor Davis' monetary reward program has also led to greater emphasis on the SAT-9. As mentioned previously, Alamitos has not yet qualified for the monetary rewards. Regardless, most of the teachers expressed concern about this reward system. Linda commented:

Not only do you have an automatic lure for cheating, in an underpaid profession, but the stress on teachers to perform, and to exclusively teach to the test, is almost undeniable... The reality of it is, teachers

start to clamor and think, "Hey, maybe I could buy a house. Maybe I could put new brakes on my car. Maybe I could travel a little bit...Maybe I could do something like other professionals who have the level of education I've had."

Mary and Nicole mentioned that at the past two staff meetings, most of the focus was on a packet of materials from the state on ethical practices in testing and preparation for the test. The packet spelled out what could and could not be used to prepare students, what could and could not be on the walls during the test, what could and could not be said or done during the actual test, and the consequences of engaging in "unethical practices." Both found the packet and the emphasis on it to be insulting. Nicole used it as an example of the "absurdity" happening in the schools as a result of high stakes testing. She commented:

It seemed to me that this was an attempt to remind teachers that, "No matter how much money is being dangled in front of your nose, you should not try to (clears throat) help your students in any way that is unethical."

In the months prior to testing, teachers were required on a daily basis to dedicate 30 minutes to an hour for test preparation. This requirement came down from the district. All teachers at each grade level were required to do the test prep during the same block of time. This was a challenge in that each teacher had a different schedule. Many had to rearrange their day to accommodate the common test preparation period. Grade level teachers were also required to turn in reports to the administrator detailing what it was they were doing during this time.

The teachers explained that during test preparation time the emphasis was on getting kids into test taking mode. Many emphasized practicing bubbling-in answers on multiple-choice answer sheets similar to those of the SAT-9. The district purchased a test preparation series called *Test Best for Test Prep* (Steck-Vaughn Company, 1998), which closely matches the format and content of the SAT-9. Harcourt, the producer and profiteer of the SAT-9, publishes this series as well. In addition to *Test Best* teachers have found, purchased, or created other materials to help students practice these tests. The teachers find it necessary to spend lots of time exposing students to the testing format as it differs greatly from their ordinary instruction, classroom assignments, and homework.

Bianca expressed her frustration:

I have caved in more to teaching to the test. [Before] I always think about ELD and

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

SDAIE strategies... I know those strategies in and out, and I've used them constantly... [But now] I've caved in a little bit to this whole pressure of the test and the scores and our APIs.

The teachers in this study have resisted teaching to the test. They have been trying to maintain what they consider to be authentic teaching by gearing learning toward the individual needs of their students, especially their ELL students. However, the emphasis on the SAT-9 and the pressure to raise scores has made it difficult to maintain their professional integrity.

Teachers' Perceptions of the SAT-9 as a Fair and Valid Measure of Teaching and Learning

When I asked the teachers, "Do you think the SAT-9 is fair?" all answered immediately and emphatically "No!" - except Mary who literally had to stop laughing first. The reasons the teachers gave to explain why they feel the SAT-9 is unfair and invalid included: (a) the mismatch between what is taught and what is tested, (b) issues of testing children who are not yet fluent in English, and (c) issues of social, cultural, and class bias. The teachers described problems with the dates the test was administered and stated that the results of the SAT-9 were of little use to them.

Mismatch between instruction and testing. The fact that the SAT-9 is not aligned with the curriculum taught in California schools has been clearly documented (Schneider & Holtzman, 1999). Ironically, the State of California spent several years developing content standards for instruction in classrooms but then purchased the SAT-9-an off-the-shelf, nationally-normed exam. Nationally-normed tests cannot be used to assess the extent to which students meet state standards, unless of course, a state aligns its standards to what is tested on the nationally-normed exam. This would be discouraged. In addition, norm-referenced tests do not inform schools and teachers as to what students are learning. They only inform schools and teachers as to how students compare to a sample of students across the nation.

For example, some of the teachers mentioned the problem of the difference between the population on which the SAT-9 was normed, and the population of the students at Alamitos. In addition to class and cultural differences, the number of ELL students in the norming population was minimal compared to the population at Alamitos. As Linda explained it:

Well it's inherently unfair because I don't think it's a valid measure based on the norming group. I believe the norming group

had a 2% second language learner population. Well, our school has probably a 90% second language learner [population]. So, consistently right off the bat, kids who don't have a full grasp of English are being asked to take a test in their second language. So, inherently, that's a problem.

In an evaluation of California's assessment program in 1998, WestEd, an independent educational policy research center hired by the state, recommended that the state stop the testing program, re-evaluate, and make the necessary changes before proceeding (Schneider & Holtzman, 1999). To date it appears that the State has ignored these recommendations.

The state of California has, however, been trying to remedy its testing system by working with Harcourt to create additional, criterion-referenced math and language arts sections of the test that are more in line with California's standards as part of the STAR Augmentation.

Added to the complexity is the fact that Beach Cities Unified has spent several years developing its own content standards. The district standards differ from the state standards. The district invested a substantial amount of time, money, and resources to develop and train teachers in the standards and created a series of its own standards-based assessments. The affect the SAT-9 has had on the district's standards and assessments will be discussed later.

Linguistic bias of testing English language learners. All of the teachers argued that the SAT-9 is inherently unfair for English language learners because they are being tested in a language in which they are not yet proficient. Bianca argued, "Of course it's not fair! Its just like if I were tested in another language, where I would be classified as the bottom of the 20th percentile."

When students are not fluent in English, the SAT-9 is not a valid measure of learning or knowledge. Lynda commented:

I think we need to ask ourselves, "What are we testing for?" Are we testing language? Then let's call it a language test, and let's make it oral. Are we testing reading? Then let's talk about what we're testing for. To try to purport that we are testing for knowledge and acquisition of skill, when we're really testing for language and reading ability is deceiving the public, it's deceiving the tax payers, and making it look like we're not providing a solid education.

Specifically, the teachers noted that ELL students simply do not yet know many of the words

used on the SAT-9. The way sentences are phrased is problematic. Mary pointed out that the test directions often use a negative in places where the students may listen for a positive (e.g., Which of the following is not an even number?). This strategy confused many of her ELL students. Some portions of the SAT-9 contain oral prompts that are read aloud to students, particularly in the math section. Teachers are only allowed to read the prompt once, which clearly disadvantages ELL students who frequently need repetition in order to comprehend.

Another problem for ELL students is that each section of the test must be completed within a specific amount of time. Hina mentioned that ELL students need more time to process in English than students who are native speakers of English. Hina speaks from experience; she had great difficulties passing one of the teacher licensing exams until she was granted extra time as a non-native speaker of English. She explained:

English being my third language, I would have to read those passages of questions a couple of times to fully understand and answer. So that extra hour on each section they gave us, that's all I needed. And I think maybe the kids are like that. They really need that processing time. And we were telling them, "OK, question number one, answer this in this amount of time, and let's go to the next one." But if you're going to give them a test, you should be able to give them enough time, and be able to answer on their own, without a time limit.

Social, cultural, and class bias. The SAT-9 was normed on a student population with less than 3% English language learners. In California, English language learners make up 25% of the student population. As mentioned above, the ELL population at Alamitos was 87%. The representation of students of low socioeconomic status was also likely much lower in the norming population than the students at Alamitos. The teachers often joke that the test was normed in Iowa, suggesting their culturally and linguistically diverse students are being measured against mainstream midwestern students. Nancy commented that the SAT-9 "might be fair in your more middle class, white Anglo Saxon Protestant household, or in a household where English is spoken fluently." Hina made a similar comment:

I think some of the questions in there are just focused on the white population, which I feel is not fair for our kids. Their lifestyles are really different from white populations, so it's really hard for them to answer those questions correctly.

The teachers gave several specific examples of ways in which test questions were culturally, socially and/or class biased. Bianca gave a personal example from her own life about the predominantly Latino neighborhood in which her parents live:

I'll go to my parents [house], and I get so mad. "Mom, I can't believe there's not a bagel place around here!" And she goes, "You need to go where the white people live over by the beach! We don't have bagels around here. Latinos don't eat bagels and cream cheese." So, it's the same thing with the test.

Hina explained why Alamitos students rarely see things beyond their immediate neighborhood:

Our kids, whatever is in their neighborhood, that's all they see. We have the beach only a few miles from us, yet half my kids have never been to the beach. That's because parents, they're working, [or] they have no car. They don't know exactly what's around us. And if they're working shifts, they usually don't have time to take their kids anywhere...In that sense, the students don't have that much experience.

Mary commented that the producers of the SAT-9 have made token attempts to make the test more sensitive to students of other cultures merely by changing the names in the stories:

The test tries to make up for that by writing stories about Carlos. You know, they put these ethnic names in these stories...like Carlos in the kitchen with his mother setting the table when the car drives up into the driveway and honks the horn. It just cracks me up that they're trying to take a test and adjust it to these kids, like, "Oh well, maybe they'll listen and pay attention because it says Carlos or because it says Juan and Maria." But that's about the only adjustment you see...and typically when they go Asian...they use Chinese names...which cracks me up because even my Asian kids will look at the names and go, "What does that say?" They have no idea...Most of the stories are very middle class [and] Caucasian.

Teachers gave several examples of the class bias in the SAT-9. Nicole mentioned that there might be a question about airplane travel, yet only two of her students have ever been to an airport. Mary mentioned a reading comprehension question on a passage about a woman who worked in the health profession. The students were required to make an inference on why the woman chose this job. The

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

answer the test creator was looking for was that she enjoyed helping people. However, Mary noticed that her few students who could actually read and comprehend the passage chose the answer, "because she needed the money." She explains her theory on why her students selected this answer:

When you bring a test like that to a socioeconomic group that meets our criteria at our school, and you ask them why people have jobs, they are going to tell you its because they need the money...And, for the kids to have to infer that it was because she had enjoyed the fact that she was caring and giving and able to take care of other people, that inference isn't there when people are at the survival level.

The fact that Alamitos students live in a low socioeconomic area also means that they are exposed to and have to deal with many things that middle and upper-class students rarely experience. Linda mentioned that last school year, out of her 20 students, 9 had fathers who were incarcerated. One of her kindergarten students told her, "I know why my daddy's in jail... Because he choked my mommy and tried to kill her." Bianca talked about a student who joined her 2nd grade class a few months after the beginning of the school year:

She cried the whole first week of school, didn't want to do anything. And I finally found out that she had been homeless. So she had only been in kindergarten. Her mother is dying. She just told me on Friday.

Bianca mentioned other challenges facing her students:

In phonics we blended [the word] *rat*. [The students said] "Oh yeah, I have rats in my room!" "Yeah, they're under the bed." "Oh, my mom can hear them in the closet!" I have another little kid who has cockroaches coming out of his backpack...Home life isn't good for them...I send them to the homework club just so that they have a clean place with a pencil in front of them to do it. A lot of these kids are dealing with more than just language issues, or being new to the country, or just even having low basic skills.

Nicole commented on how the problems her students face on a daily basis affect their ability to learn:

I've had children who have come in whose fathers had been shot for their wallet, [or] who've been beaten black and blue. Children whose parents have gone off the deep end ...I admit, there's alcoholism all over. There's

child abuse all over. I'm not saying [higher socioeconomic schools] don't deal with some of these things, but they don't have the added emotional strain of children who are so needy.

Nicole argues that these factors and others that students deal with cannot be accounted for in a standardized test:

There's pressure from below from the students who are looking at you going, "I can't learn what you're teaching me because my mother's in the hospital. My brother got taken to jail. We were under the beds all night long because there was gunfire." I don't think they can norm that for a test. Sorry. I think they should norm this on sleep deprived, terrified, underfed children.

Nicole mentioned the well-known statistic, "The higher the SES (socioeconomic status) the better the kids do [on standardized tests]." One issue also common in low socioeconomic neighborhoods is the low level of education attained by the parents. Linda mentioned, "The mother's educational level...is a widely known and widely accepted indicator of student ability." The majority of students' parents at Alamitos do not speak English fluently. In fact, there are many parents who are not literate in either their native language or English. This severely limits their ability to help their children, as Hina explains:

Our parents, even though they want to help their kids, their English is not so high, so its very hard for them to help...Even if they're educated in Spanish, it's very hard for them to help them in English.

Linda gave an example of how the differences in socioeconomic status and education can disadvantage students in terms of knowledge gained outside of school:

And we also know that parents who aren't literate, for whatever reason, have less to give their children in terms of materials in the home, opportunities to meet texts; therefore they are always at a disadvantage when it comes to doing this kind of work. Their resources at home are less, and consequently, the test is inherently unfair because...it's affected too much by outside variables over which we have very little immediate control...Take a first grader who is being subjected to the SAT-9...if his home environment hasn't prepared him, you're stacking him up against a kid whose mother may have gone to the library every week, and may have gotten him educational videos and programs for him. Therefore...at

home...he may not have acquired that kind of book learning and school knowledge that's being asked for and tested on the test.

As Linda pointed out, many children begin school at Alamitos with very little experience with print. When I taught kindergarten at Alamitos, I discovered very quickly that many of the students never held a pencil, or had someone read a book to them. During home visits, I noticed very few books or reading materials in the home. Many Alamitos students have a clear disadvantage in that they are being held to the same standard as middle and upper class children who likely have parents who read, who buy them books, who take them to the library, and who read them bedtime stories.

Much of the SAT-9, even in grades 1 and 2, requires that the students read instructions, passages, and answers on their own. The teachers argue that if students are not capable of reading the exam, then it is not a valid measure of what they know, especially in areas other than reading comprehension, such as word problems in the math section. Thus, the SAT-9 becomes a futile endeavor for some students in bubbling in answers at random. The results are completely meaningless. When I was still teaching at the school, a 1st grade teacher told me about a strategy one of her students used to mark his answers. He pointed to each bubble while chanting, "Batman, Batman, fly away," and marked the answer wherever his pencil landed.

As shown in Vignette 1, students were trained at school-wide assemblies and by their teachers to mark an answer for every question, even if they had to guess. Bianca mentioned having her aide stand next to a student who had recently arrived in the United States, mainly to encourage the student to mark random guesses for every question on the test. Since there is no penalty for wrong answers, students had a 25% chance of getting each question right if they guessed, but no chance if they leave a question blank. The school could raise composite test scores with just this strategy.

Issue of testing dates in year round schools. The fact that Alamitos is a year round school has created some problems in terms of the dates during which the SAT-9 is administered. Up until the 1999-2000 school year, the state had designated a specific time period in which all schools had to administer the SAT-9 (usually around May). The strict testing time period meant that students on certain tracks had between 30 to 60 days less instruction than students in traditional calendar schools, which are typically in higher socioeconomic areas. So, not only were students disadvantaged on the SAT-9 in terms of the linguistic, cultural, and class issues discussed earlier, they were also

disadvantaged by having to take the exam with significantly fewer days of instruction than more advantaged students. Harcourt was supposed to adjust for this statistically, but, in addition to other mistakes, forgot. They forgot two years in a row. The scores year-round schools posted were lower than they should have been. This made the headlines (Colvin, 1999; Djurklou, 1999) and the California Department of Education fined Harcourt \$1.1 million dollars. Adjustments or not, Alamitos teachers still had the added pressure of trying to cover everything on the exam in a shorter amount of time.

Another issue of SAT-9 testing in year-round schools is the fact that hundreds of students take the test immediately after returning from a month-long vacation. At Alamitos, in 2000, these students were on Track B. They were off in April and began testing the first and second week in May. Track B teachers had to prepare their students the best they could and hope the students would retain most of it over the vacation. Track B contains the largest number of Asian students. This may help explain why Alamitos failed to meet its API growth target for this subgroup that year. When the testing dates were changed in 2001, the majority of students who were on vacation immediately prior to testing were Latino(a). Incidentally, Alamitos met the growth rate for Asian students in 2001, but not for Latino(a) students.

Usefulness of SAT-9 results to improving instruction. I asked teachers, "When do you get the results of the SAT-9?" Mary laughed and said, "Typically either right before the school year ends, or within a month and a half of the next school year." Bianca also laughed and said, "After you've already started a new year with a new group of kids!" After all the hard work the teachers put into preparing their students for the SAT-9, the results come back so late that they are of little value in helping the teacher pinpoint the needs of the students. The results are of little help to the students' teachers the following year as well. Mary explains, "By the time that I get them [the results], I've already figured out which students need the extra help or need remediation."

One problem with the results is that they are so generic they do not provide information on any specific skills or problems the children were unable to solve. This is a given; however, when norm-referenced tests are used to assess the extent to which students meet state standards. Categories of deficiency are too ambiguous to inform remediation efforts. Reports only provide scores for sections and subsections.

At a staff meeting at the beginning of the 1999-2000 school year, for example, teachers were given test results showing only the average for each

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

section across grade levels. All 2nd grade teachers had only the grade level average with which to work. After viewing these averages we met in grade level groups, discussed the sections in which our students scored low, and brainstormed ways to focus on these areas during the year to raise the scores, not to mention to improve the learning in these areas. To assist us with this task, the district provided copies of the previous years' SAT-9 exam. We were able to view the test and discuss it, but we were not allowed to take any notes or write anything down. Regardless, this approach was not very effective in helping us plan our curriculum. First, because all classes were averaged together, there was no way for us to know how well our own class did on certain sections of the exam. Second, we were looking at scores of students who were no longer in our classes. Third, there was no way to tell which particular problems gave students difficulty. The only "benefit" that came from the exercise was that teachers were exposed to what the test looked like, what was important on the test, and what would be important in their teaching the following year. The test results were not and could not be used to pinpoint the educational needs of individual children.

Effects of the SAT-9 on the Curriculum

The teachers reported that the district's emphasis on the SAT-9 has resulted in drastic changes to the curriculum. These changes include (a) changes to district standards and assessments, (b) the adoption of new curricular materials in math and language arts, and (c) the de-emphasis or elimination of content areas and other activities not related to the SAT-9. The teachers did not consider these to be positive changes.

District standards and assessments. The district has spent several years creating its own content standards and authentic standards-based assessments. The idea behind standards-based reform is that the standards drive the curriculum and assessments are used to determine how well students meet the standards. However, teachers reported that the SAT-9 is now driving the standards.

An analysis of the district's standards revealed that the standards were changed to include concepts covered on the SAT-9. A page from the district's revised math standards is shown in Figure 1. The lines indicate the 16 new items that were added on this page for K, 1st and 2nd grades to match the SAT-9.

An analysis of district assessment instruments reveals that they too have been revised to be more reflective of both the content and format of the SAT-9. For example, the district has an authentic reading assessment program based on identified

benchmark books that students are expected to read at certain points in their schooling (e.g., end of kindergarten, emergent, end of 1st, middle of grade 2, end of grade 2, etc.) about which they must answer comprehension questions. In response to the SAT-9, the district revised the assessment program by pushing the benchmark books down one level, (e.g., middle of grade 2 became end of grade 1) in order to match the level of the reading difficulty found on the SAT-9. The comprehension questions were also changed so that at least half were multiple-choice and were of the same bubble-in format as the SAT-9.

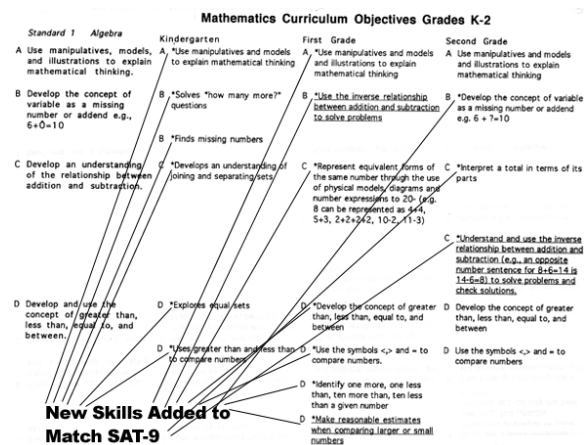


Figure 1. Sample page from the district's Revised Math Standards showing skills added to match the SAT-9.

Math. Most of the teachers mentioned the adoption of *Math Steps* (produced by Houghton Mifflin) which is essentially a series of worksheets emphasizing paper and pencil practice. Mary explained that the series is "aligned with the California content standards...so they thought it would align more with what's on the SAT-9." The publisher's description supports Mary's assertion. An advertisement for the series describes it as "carefully sequenced lessons...that prepare students for standardized tests" (from www.schooldirect.com). On every page, there is a special section called "Test Prep" where the practice item appears in the same bubble-in format as the SAT-9. This new series was in stark contrast to the previous series, *Anytime Math* (ironically also produced by Harcourt) that emphasized hands-on, exploratory, self-discovery math and de-emphasized paper and pencil tasks.

In addition, the district created a number of its own required math materials and assessments, such as open-ended math (OEM) prompts. The district is also requiring students to complete a

number of "learning records," which are entries into a math journal in a style similar to OEMs. The idea is that students will have a notebook to refer to with which they can study for the SAT-9. While teachers think the concept is good, they feel in practice it is an unproductive use of time, especially for ELL students. Hina commented, "Just doing one learning record takes up the whole 45 minutes, and we really don't get any other teaching in there!" Bianca noted that for ELL students "It ends up being a huge copy session."

The large number of new math materials, assessments, and procedures have left many teachers in the district confused. As Mary explained: "What you get is this mish mash, and you still have teachers trying to figure out, 'well, what exactly do you want me to teach?'" I attended a training meeting with Mary where the district's math curriculum leaders attempted to make sense of the myriad of new materials and requirements. We were expected to take the information back to Alamitos and explain it to the other teachers. One of the items given to us at the training meeting was a math correlation guide, which essentially listed all of the math skills covered on the SAT-9 with correlations to exact page and problems numbers in *Math Steps*, *Anytime Math*, *Test Best for Test Prep*, open-ended math prompts, standards, assessments, and other math materials.

The following school year this evolved into a comprehensive pacing guide designed to ensure that students would be exposed to all of the math concepts before the SAT-9 was administered. The teachers found the pacing guide to be completely unrealistic and, again, were perplexed that it was based on the traditional calendar year yet was being mandated for year-round schools. The pacing guide required teachers to squeeze into six months what they would normally teach over a nine-month period. Adding to the stress, teachers at Alamitos were expected to teach more material, more quickly, and in less time than was the case in higher SES, traditional schools.

Nancy, expressing her concerns about this, gave a specific example: "It had things such as teaching money in two days. I'm terribly sorry. Maybe on the Eastside, where students have their own bank accounts."

Hina expressed her frustration about losing control over the decision of what and when to teach math concepts to her students: "This [the pacing guide] is so ridiculous! Every class is different...you really have to see where your own class is and go from there." It appears the district is concerned that if teachers slow down to make sure their students understand before moving on, then the higher students will not get exposed to other concepts and will get them wrong on the test. The result is a

sacrifice of the learning of lower students in order to ensure that higher students get everything they need to raise the school's average math scores.

Language Arts. Teachers reported that the most dramatic changes to the curriculum came in language arts. Prior to the third year of SAT-9 testing, the district's literacy programs were based on whole language philosophies. Teachers emphasized shared and guided reading and used books which provided support to emergent and early readers (Fountas & Pinnell, 1996). Sound-symbol relationships were taught in the context of these books, and through activities such as shared and interactive writing (McCarrier, Pinnell, & Fountas, 1999). The district's literacy leaders and trainers discouraged the teaching of phonics out of context.

Two months into the school year, however, the district adopted the *Open Court* phonics program for grades K-3⁵, provided emergency training for all teachers, and required the immediate implementation of the program. The program takes approximately two hours per instructional day, is highly scripted, and prescribes exactly what it is the teacher is to do, say, and write on the board. Specific dictation drills, worksheets, and readers for each lesson are also included. *Open Court* is a one-size-fits-all program. All students, regardless of their levels of performance, are submersed in the exact same lessons. Ironically, there was little or no time left for reading.

The district reassigned a large number of literacy and ELD specialists as "*Open Court* Coaches," to ensure strict and exact implementation of the program. The coaches, who came to be known as the "*Open Court* Police," patrolled classrooms and made sure teachers were following the script properly.

The teachers insisted the decision to adopt *Open Court* was because of the SAT-9. At a meeting I attended, a top district official confirmed this was true. He acknowledged he made the decision to adopt *Open Court* because of its success helping to raise SAT-9 scores in Texas and the Inglewood and Sacramento school districts in California (for counterpoints to this claim, see Coles, 2000; Moustafa & Land, 2001; Taylor, 1998). There was also substantial pressure from the state and a member of the State Board of Education to adopt the program (Lapp & Flood, 2000, Fall).

The teachers resent the program because they are no longer allowed to use their professional judgment, experience, and expertise to design individualized literacy instruction. They described the program as "too boring" for students who can already read, and "too difficult" for students still struggling.

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

Their major concern is the large amount of time taken up by *Open Court*. Mary commented, "*Open Court* takes a great chunk out of your teaching day, [and] just specifically for a couple of sections on the test."

Bianca expressed her concern about the appropriateness of *Open Court* for ELL students. She found her students did not know many of the words used in the phonics drills. It takes her even longer to present each lesson, taking more time from her regular instruction.

Hina noted she has significantly less time to meet with students in small reading groups. She complained about the black-and-white plain paper books that are used to emphasize the phonics elements taught in the lesson. She said, "They are learning how to sound out and blend, but there's no comprehension in the books. They make no sense! And the language used in there goes over our kids' heads!"

The reality, as the teachers have explained, is that *Open Court* became the language arts program. *Open Court* takes up about one-third of the teaching day although it was supposed to be a supplemental resource. There is very little time left for shared reading, guided reading, shared writing, interactive writing, or to just read books aloud to the students. The amount of time to work individually with students in small groups also greatly decreased.

In my classroom prior to the adoption of *Open Court*, I was able to meet with each reading group for 20 minutes and had time to read with every student a book matched to their reading ability. Once *Open Court* was adopted, these small reading groups became 15-minute "workshops." I was required to use *Open Court* materials instead.

For the first time in many years, the district also adopted a basal spelling program. Before, teachers were afforded their choice on how they would teach spelling and were trained to observe and analyze student reading and writing to inform spelling instruction. Because one of the sub-tests on the SAT-9 covers spelling, once again teachers were sent to mandatory training meetings to be trained in a new program. Teachers were instructed to give students the same spelling lessons and spelling words from the basal, regardless of their spelling abilities.

The other content areas: Social studies, science, ELD, art, music, and PE. Between *Open Court*, math, ELD, writing, and mandatory test preparation time, teachers reported that there is no time left for anything else. Nicole asked, "Where do you fit it all in?" Figure 2 shows a time allotment and sample schedule given to teachers during an *Open Court* Training Session. The boxed items were to be

given top priority. If assemblies, photos, field-trips or other interruptions occurred during the day, teachers were still expected to provide instruction for the full amount of the stated time for those items. When one considers that *Open Court*, *Math Steps*, and the basal spelling program were adopted specifically to raise SAT-9 scores, this schedule clearly shows that teachers were given no choice but to teach to the test, and that teaching to the test was an all-day endeavor.

Sample Daily Schedule Implementing Open Court Grade 1-3 300 minute schedule	
Items listed in the boxes are required on a daily basis regardless of how a daily schedule is impacted. (e.g., assemblies, early dismissal etc).	
8:00-9:00 60 min.	L.A. Content Standards 4 & 6 Phonemic Awareness & Phonics -Introduction and Review of Sound/Spelling Cards -Blending -Dictation -Workbook practice pages
9:00-10:00 60 min.	L.A. Content Standards 4, 5, 6, 7 Workshop-Balanced Literacy Program (see attached workshop page) • Literacy Centers (Read To) -Words Their Way/Words/4 Corners -Super Reader /Self Recording/Buddy Read/DEAR (Ind. Reading) -Small Group Instruction/Individual Practice (Instruction) -SAT 9 Skills Practice (Instruction/Ind. Practice)
10:00-10:15	15 min. RECESS
10:15-10:30 15 min.	L.A. Content Standards 5 & 6 -Comprehension/Content/SAT 9 Direct Instruction and Practice -Expository Text/Thematic Instruction -SAT 9 Practice-strategies and skills
10:30-11:00 30 min.	E.L.D. Content Standards E.L.D.-Into English
11:00-11:30 30 min.	L.A. Content Standards 5 & 6 -Comprehension/Content/SAT 9 Direct Instruction and Practice -Expository Text/Thematic Instruction -SAT 9 Practice-strategies and skill
11:30-12:05	35 min. LUNCH
12:05-12:15	Content Standard 7 DEAR/Buddy Reading/Super Reader
12:15-1:15 60 min.	Math Content Standards Math-test prep (strategies and skills)
1:15-1:55 40 min.	L.A. Content Standards 4, 9, 10, 11, 12 Writing/Spelling
1:55-2:05	10 min. Test Prep-format
2:00	20 min. P. E. (200 minutes every two weeks, which happened at grade levels)

Figure 2. Time allotment/sample schedule.

District level administrators told teachers they were not to teach art or allow students to draw pictures to accompany their writing as part of an ELD lesson. The rationale for this, Linda explains is, "because we're not going to be drawing a picture on the SAT-9." Teachers reported music is down to five lessons a year with the music teacher. Physical Education (PE) rarely happens before testing even though law mandates 100 minutes of PE instruction per week. Mary mentioned that at a staff meeting this year, teachers were given a thick packet containing state PE standards, activity cards, and a comprehensive teacher's guide. She said when the materials were passed out, the message from the administrators was, "I don't know when you're going to teach this...or what equipment you are going to use...but in your spare time, you can look through it."

Linda added that some kindergarten teachers at Alamitos have even eliminated recess.

In addition, the teachers lamented the fact that there is not even time left in the day to teach social studies and science, subjects they know are just as important as reading and math. Many of the teachers mentioned that the district adopted a new social studies program this year. Teachers were even sent to a training seminar on it. Mary said the program "is wonderful, and it aligns with our ELD curriculum," but expressed frustration about having no time to use it: "I literally have not opened it with the students. The kids haven't even seen it." Bianca made a similar comment: "The brand new social studies [program] is fabulous! But have I opened it with my students? No!"

Science instruction has always been one of Nancy's strengths, and her emphasis on science has been one of the defining characteristics for her Excel program. However, with all the changes related to the SAT-9, she laments that science and social studies instruction in her classroom has been greatly reduced as well: "It's miniscule, absolutely miniscule in comparison." Bianca also described her struggle to provide social studies and science instruction:

The only way I can do social studies or science, for that matter, is ELD. And it's not truly like a well-rounded social studies unit. It's just kind of like, we touch on it a little...It's not a real outstanding unit you can really get into.

Additionally, teachers are finding it difficult to provide daily ELD for their ELL students also required by State law. Teachers are finding that *Open Court*, District assessments and SAT-9 preparation are increasingly encroaching on the time previously devoted to ELD instruction as well. Bianca explains, "ELD used to be my favorite time of the day to teach. I love ELD. Well, sometimes if we have to do an open-ended math prompt, or SAT-9 prompt, I'll skip over ELD." Mary said:

I like teaching the ELD curriculum, even though I don't get to it as much as I would like to. Because of all the priorities of the day...I don't hit ELD on a daily basis. And when I do, I sometimes have to tie it in with my spelling.

The teachers commented that in addition, there is now less time to develop students' oral language throughout the day. Teachers have found there is no time to engage the students in class discussions, one of the most effective techniques to develop the oral language skills of ELL students. Linda cracked, "It's like we say to the children, 'Sorry, I can't listen to your story about your puppy because I have to read you this book about a dog.' It

was often in these class discussions that teachers could discuss openly with students some of the challenges they faced in their homes and neighborhoods. The teachers feel frustrated. Because of the emphasis on the SAT-9, they have less time to meet the social needs of their students.

Issues of educational equity. The fact that students are only being taught subjects which are tested on the SAT-9, to the near exclusion of social studies, science, art, music, and PE, raises some serious questions of educational equity, particularly for ELL students in inner-city schools. I asked teachers if schools in the higher socioeconomic areas were still able to teach these subjects. Linda and Mary both have children at elementary schools on the eastside (higher socioeconomic area) of the district, and confirmed that eastside schools were still teaching social studies, science, art, music and PE. Mary expressed her disgust with this situation:

What upsets me about that is the law states that there should be equal opportunity in education. And my belief is that our kids, because they are who they are and where they come from, are being punished for that, and not being allowed an equal opportunity in education.

The first grade teachers are especially frustrated about the emphasis on the SAT-9 and the narrowing of the curriculum. The state does not require 1st graders to be tested. The decision to test 1st graders was a district decision. Many teachers speculated that the district wants the 1st grade students to experience the SAT-9 one-year before the SAT-9 counts in the API. This way, when students are 2nd graders, they will be more familiar with the test.

To mediate the situation, all twelve 1st grade teachers and four 2nd grade teachers decided to voluntarily extend their instructional day by one-half hour, without pay, to provide students with the full curriculum they would have received had their not been so much emphasis on the SAT-9. Hina explains why they made this decision:

This year, we...extended our day a half hour more. And this is exclusively to do science and social studies...We think it's very important for our students to learn other subjects besides *Open Court* and math...because in upper grades, their literature, all that is based on social studies, and science and things like that. And if they don't get that base from the beginning [in] 1st [and] 2nd grade, they're going to have a very hard time understanding the literature in the upper grades...There was no room for social studies, science, so that's when we

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

decided to extend our day a half hour...But this is a time for us. With that half hour, we can teach whatever we want, and especially in social studies and science and stuff, and not have to worry about, "OK, this is what we have to do." It's our own time, and we pick what we want to do.

Bianca explains the feelings of teachers that led up to this decision:

There are a lot of teachers who feel, "I'm not meeting the needs of these kids because I'm teaching to this test, which is going way beyond what they really need...I'm not doing the best that I could for my kids." [So] they're saying, "OK, let's move the day to 2:15, without pay."

Many other teachers, however, are very concerned about the precedence this decision is setting. Nicole expresses the concerns of many teachers:

We are paid at a very low level considering the amount of education that [we] have, and the amount of work that we are expected to do is escalating. Testing...is generating a massive amount of paperwork. So, we're doing more, getting paid the same, and then there are teachers who are saying, "Oh well, we'll work for another three weeks for free," which is what the time after school multiplied over the year works out to be. And eventually, as people know, common practice becomes the rule. So it will be, "Well, these people have extended their day, so we will extend everybody's day, without pay."

Nicole's words proved to be prophetic. A few months after our interview, Linda wrote to inform me that a vote was taken at a staff meeting, and the majority of teachers elected to extend the school day a half-hour, without pay. She noted how pleased the principal was and said that he congratulated the staff for their professionalism.

Teachers already spend many hours beyond the required work day grading papers, organizing their classrooms, and preparing lessons and materials for the next day. But to voluntarily extend the instructional day with students raises a number of issues, especially in terms of the contract between the teachers' union and the district. But obviously, these teachers were not thinking about themselves. As Linda explains, "This proves what kind of character teachers really have. In other words, these teachers believe so strongly that kids need this, that they are willing [to provide it] essentially for free." However, it was clearly a divisive issue at the school, and many

of the other teachers felt the pressure or guilt to extend their day too. There is also the risk that someday the extended day will also be consumed by SAT-9 preparation.

Effects of the SAT-9 on Teachers and Students

As evidenced in the examples above, the teachers are feeling stressed and overwhelmed because of the great emphasis on the SAT-9 and the pressure they feel to raise SAT-9 scores. The teachers are feeling disempowered as they are being forced to implement one-size-fits-all programs that they know are inappropriate for their students. They are growing tired of being compared with teachers of middle class and native English-speaking children and absorbing the blame for their students' low scores. They are insulted by monetary rewards and wonder why more monies are being given to advantaged schools instead of the schools that need equitable resources.

The teachers had a lot to say about the effects this is having on themselves and others at Alamitos. Linda said:

I think people are walking around with almost a palpable, visible hang-dog feeling. The feeling that, "I no longer matter. Why am I educated?"...And I've seen a lot of teachers who want to quit the profession, who are saying, "This is not what I signed on for."

Nicole noted a conversation she had with 3rd grade teacher:

I asked a pointed question to a teacher at my school. I asked, "Do you know anybody on this campus who is happy? Any teacher who is happy?" "No." It came back that fast (snaps fingers). "Do I? No." Large staff, nobody is happy. I'm watching the teachers become very frustrated by watching good teachers leave.

Bianca mentioned a recent experience with one of her 2nd grade colleagues:

Sonia's like, "There's so much, I'm so overwhelmed, I'm so exhausted, I'm so stressed! But, I'm such a bad teacher." [After a staff] meeting she just started bawling. I mean, bawling! Just from the stress, and just from feeling, "I'm not good. I'm not a good teacher." And Jenny and I were just consoling her saying, "You are one of the best teachers we have."

Nicole made a comment that is very telling of the feelings of teachers:

The most pathetic thing is that up until about two years ago, I counseled young people, "Come into teaching. It is a wonderful profession." Now I counsel them to find

something else because this is not the profession I would choose for myself.

Kindergarten is the only grade level that is not required to administer the SAT-9. However, from Linda's many comments above, it is clear that the emphasis on the SAT-9 has had a dramatic effect on the kindergarten program, teachers, and students as well. Linda mentioned the stress she sees in her teaching partner and the effects tests have on her students:

I see my partner for example. "Hurry up! Sit down! Write! Hurry! You've got to go. This is all being done to kindergarten kids, where there is absolutely no consideration given to their fatigue. In fact, words like this have been used: "You need to build their stamina for longer [reading and] writing!" This is coming from our administrator...So now school has become a survival ordeal...It breaks my spirit to see little kids, who should be imbued with nothing but enthusiasm for learning, to take that enthusiasm, and kind of corrupt it by forcing and channeling it into these little narrow definitions of what learning is.

The teachers reported many examples of how the emphasis on the SAT-9 is having an effect on the students. Mary told me:

I've had kids break down and cry, when I've administered a test. And it just breaks my heart because all you can say is, "Just try your best." And then the other end of the spectrum is the ones who just don't care, and they figure just fill in any bubble and get it over with, because as soon as you close your book you don't have to do anything else with it. So, that's what you get-apathy; but how pathetic is that that 2nd graders should have to experience that?

Vignette 2 below depicts an event described by Bianca, which took place in one of the 2nd grade classrooms at Alamitos:

"I Know How to Read a Word"

The 2nd grade students have their test booklets open on their desks and sit quietly as Mrs. Sanchez guides them through the sample problem of the reading comprehension section of the SAT-9. Mrs. Sanchez finishes reading the text from the teacher's test administration booklet, looks up and asks, "Are there any questions?" Her question is met by mostly blank stares. She then continues reading from the test booklet, "You have 50 minutes to complete this section. You may begin." A few students pick up their pencils and begin to read the

first passage. A few start bubbling in answers without bothering to read the passage first. Jorge, a Spanish-speaking ELL student, looks around, looks at his test booklet, looks around again, and finally looks over at his teacher who is sitting at her desk keeping a watchful eye on the students. As they make eye contact, Jorge leaves his desk and walks over towards her. In a sad voice he says to her in Spanish, "Mrs. Sanchez, I can't read. I can't read it!" Mrs. Sanchez gives him a sympathetic look. "I know Jorge. Just do the best you can." Jorge returns to his seat. He looks intently at the test booklet for several minutes. He then stands up and walks back to the teacher's desk, holding the test booklet. He says softly, "Mrs. Sanchez, I know how to read a word." He points to and reads the word, "the." He was so sad. Mrs. Sanchez, stroked his hair gently, wishing she did not have to make him go through this. "That's right Jorge," she said gently. She then sent him back to his desk to finish the test.

After describing this event, Bianca remarked, "You know, this does bring them down." Hina has had 1st grade students break down and cry. She described a student who had arrived in the United States just two months before SAT-9 testing:

One of my students cried the day before the test. We had been practicing for it. [She] recently came from Mexico. She had no English at all, very few words, here and there. And she was with me, I think two months...and she had to take the test. I mean, she had a hard time just when we did practice...She went home and she was crying to her mom, and her mom came and told me the next day. [The child] was so overwhelmed, the poor thing. It was really hard for her.

The emphasis on the SAT-9 and raising test scores can hamper the relationships between teachers and students. The counselor assigned the "lowest of the low" ELL Spanish-speaking 2nd grade students to Bianca, even though she was no longer allowed to teach in Spanish. In fact, the counselor transferred some of her higher students to other classrooms to make room for these low students. Given the fact that she was purposely given the lowest students, she resented the fact that she was being held to the same expectations as the other teachers to raise SAT-9 scores. Bianca described the constant struggle of trying to balance the needs of her students with the pressure to teach to a test she knew was well beyond the capabilities of her students. Linda commented on

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

how the SAT-9 can change how teachers view their students:

But if you are too stressed, that stress not only affects your health, but it affects your relationship with your students. And you start to look at your students as high scorers or low scorers, and that starts to affect your relationship. And you start to, maybe not necessarily want that kid in your class next year, and you find out that the little low kid is coming, and you kind of, don't want him. Then what does that do to that child's self-esteem when, because of no fault of his own, he struggles to do well on tests?

Conclusion

I asked the teachers, "Has the SAT-9 made you a better teacher?" and "Has the SAT-9 improved teaching and learning at your school?" Of course, "better" is subjective, but I left it up to the teachers to see how they would interpret the question. SAT-9 scores have gone up at Alamitos each year, so one could argue that the teachers are doing a better job and that teaching and learning has improved at the school. However, these proved to be emotional questions for the teachers. The answer to both questions was very clear—"No."

The five elementary school teachers interviewed in this study were extremely knowledgeable and provided substantial information needed to answer the research questions. They believe the SAT-9 is not a fair and valid measure of teaching and learning for several reasons, including the mismatch between the curriculum and the test, the differences in the school population and the norming population, the testing of students before they are fully proficient in English, cultural and class bias, the testing of some students who have had significantly fewer days of instruction than other students, and testing some students immediately after a month-long vacation.

The SAT-9 has affected the curriculum provided to students as major changes have been made to teach to the test. These changes include: revising district standards and assessments to more closely match the content and format of the SAT-9; adoption of *Math Steps*, *Open Court*, *Test Best for Test Prep*, and a basal spelling program; an increase in the number of required Open Ended Math (OEM) assessments and learning records; and mandatory time allotments for phonics, math, and direct SAT-9 test preparation. As the SAT-9 tests only language arts and math, there is no room for science, social studies, PE, music and art. Even English Language Development (ELD), which is required by State law as essential instruction for ELL students, is beginning

to be cut back. Teachers have little time during the day to develop the oral language skills of their students. The emphasis on the SAT-9 is so great that the district decided to also require the testing of 1st graders, even though the state does not. In response, many teachers have voluntarily extended their day without pay to provide their students instruction in the content areas not assessed on the SAT-9. These teachers feel their students are equally entitled to the educational opportunities in which their wealthier peers continue to engage in their schools.

The SAT-9 is having harmful effects on the teachers and students. The teachers are stressed and overwhelmed by all the curricular changes and pressure to teach to the test and raise scores. They feel they are disempowered as professionals, and are no longer able to make decisions on how to best meet the needs of their students. They are tired of being compared to higher socioeconomic schools with significantly smaller numbers of ELL students and being blamed for their students' low scores. They experience additional stress when helping students in poverty deal with the problems within their families and neighborhood. They are insulted when monetary awards are disbursed to schools and teachers of more privileged students. They are frustrated when they watch good teachers leave the school, and sometimes teaching in general.

The students feel the stress as well. Students often break down and cry because of test-related anxiety. Some are developing apathetic test-taking behaviors in the primary grades because they deem it pointless to read a test question before bubbling in a wrong answer. The pressure on the teachers to raise scores is changing their relationships with their students as well. Teachers want to see their students as more than candidates for high and low test scores, but this is difficult when pressures are placed on teachers to post gains year-to-year and meet academic targets set outside of the classroom.

These teachers strongly feel that the SAT-9 is not improving teaching and learning. The SAT-9 is having the opposite effect. The teachers maintained that if they could change anything about the current situation, they would abolish the SAT-9 and other high stakes tests. This does not mean they are against holding students, teachers, and schools accountable. All of the teachers argued for the use of authentic assessments instead. Authentic assessments show where students are at the beginning of the year and illustrate academic growth until the end of the school year. If all children are different and progress at different rates, they argue that it is much more reasonable to hold teachers accountable for each

individual's growth per year instead of helping all students pass an arbitrary standard.

Bianca adequately summed up the feelings of the teachers in this study and many other teachers who experience the effects of standardized testing:

I wouldn't have any kind of standardized testing. Let's just focus on the assessments the district has now, which up to now have worked, and which we've seen have results, and so we are more able to teach to each student's needs, and not teach to a whole group like what we're doing with a standardized test. I think that's the whole flaw with standardized testing, teachers are just, whether they like it or not, they're teaching to a whole group, and that's not the way. We all know that each child learns differently. How can you be a good teacher when you're teaching the whole group and not teaching to each student's needs?

At Alamitos Elementary School, in the Beach Cities School District, the SAT-9 is clearly not achieving its stated purpose-improve teaching and learning in the schools. Rather, it is hampering authentic teaching and learning, and harming teachers and students. In addition, it is having a differential impact on the education of ELL students in low socioeconomic neighborhoods, creating a situation where the linguistic and cultural needs of students are being ignored, and denying these students an equal opportunity to learn.

The findings of this study are consistent with the studies, reports, and position statements described in the literature. These data from the teachers, along with the data from the document analyses and my first-hand observations, provide evidence that high stakes testing is indeed having harmful effects on teachers, students, and the curriculum. While this small study focused on the use of the SAT-9 in a single school in California, the fact that similar conclusions were drawn in Smith et al.'s (1989) study of elementary schools in Arizona and McNeil's (McNeil, 2000) study of high schools in Texas where different yet similar tests were used, provides compelling evidence that there are serious flaws in using high stakes tests to improve school performance. Indeed, the findings of this study support McNeil's assertion that high stakes tests are a "contradiction," Swope & Miner's (2000) assertion that "the testing craze won't fix our schools" and Kohn's (2000) assertion that high stakes tests are actually "ruining the schools."

Notes

[1] Proposition 227 was designed to bring an end to bilingual education in California. However, some

bilingual programs have continued through a limited waiver process.

[2] All names of places and people have been changed to maintain the confidentiality of the participants and their colleagues.

[3] The bilingual program at Alamitos was dismantled after the passage of Proposition 227.

[4] I developed a set of 48 codes for each of these themes, and then inserted the codes into their appropriate positions in the interview transcripts (Merriam, 1998; Stake, 1995). For example, the code "TTTT," stood for "teaching to the test" and was inserted into sections of the transcripts where teachers described pressures or practices of teaching to the test. Separate documents were created for each code, transcript segments were copied and pasted accordingly, and the documents were analyzed and organized per research question. Since I focused on the experiences of the teachers, I relied heavily on their direct quotes to portray precisely what it was the teachers were attempting to convey (see Emerson, Fretz, & Shaw, 1995).

[5] The District adopted only the phonics component of the program.

References

- American Educational Research Association. (2000). *Position statement concerning high-stakes testing in PreK-12 education*. Author. Retrieved from the World Wide Web, January 26, 2001, <http://www.aera.net/about/policy/stakes.htm>.
- Amrein, A. L. & Berliner, D. C. (2002, March 28). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved March 30, 2002 from <http://epaa.asu.edu/epaa/v10n18/>.
- Amrein, A. L., Berliner, D. C., & Biddle, B. J. (2002, Unpublished manuscript). In pursuit of better schools: What research says about high-stakes testing and student learning. Research funded and published by the Rockefeller Foundation: New York.
- Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education: An introduction to theory and methods* (2nd ed.). Boston, MA: Allyn and Bacon.
- Butler, Y. G., Orr, J. E., Gutierrez, M. B., & Hakuta, K. (2000, Winter & Spring). Inadequate conclusions from an inadequate assessment: What can SAT-9 scores tell us about the impact of Proposition 227 in California? *Bilingual Research Journal*, 24(1 & 2), 141-154.
- Coles, G. (2000). *Misreading reading: The bad science that hurts children*. Portsmouth, NH: Heinemann.

The Effects of High Stakes Testing in an Inner-City Elementary School: The Curriculum, the Teachers, and the English Language Learners

- Colvin, R. L. (1999, July 13). New errors delay Stanford 9 results. *Los Angeles Times*, pp. A3.
- Djurklou, A. (1999, July 7). Second mistake in test scores. *Long Beach Press Telegram*, pp. A1.
- Emerson, R. M., Fretz, R. I., & Shaw, L. L. (1995). *Writing ethnographic fieldnotes*. Chicago: University of Chicago Press.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: MacMillan.
- Fountas, I. C., & Pinnell, G. S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Garcia, P. A. (2001). Use with caution: California's Academic Performance Index is misleading for schools with large numbers of English learners. *NABE News*, 24(5), 11-12, 32.
- Gould, S. J. (1996). *The mismeasure of man*. New York, NY: W. W. Norton Company, Inc.
- Grisham, D. L. (2001). High-stakes testing in our schools: A new report from California. *Reading Online*, 51, Retrieved from the World Wide Web January 26, 2002, <http://www.readingonline.org/editorial/july2001/index.html>
- Grissmer, D., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving student achievement: What NAEP test scores tell us*. Santa Monica, CA: RAND Corporation [Online]. Available: <http://www.rand.org/publications/MR/MR924/>
- Haney, W. (2000, August). The myth of the Texas miracle in education. *Education Policy Analysis Archives*, 8(41), Retrieved April 1, 2001 from the World Wide Web: <http://epaa.asu.edu/epaa/v8n41/>
- Heubert, J. P. (2001). *High-stakes testing: Opportunities and risks for students of color, English-language learners, and students with disabilities*. In M. Pines (Ed.), *The 21st Century challenge: Moving the youth agenda forward*. Baltimore, MD: Sar Levitan Center for Social Policy Studies, Johns Hopkins University Press.
- Heubert, J.P. & Hauser, R.M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Klein, S. P., Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2000). What do test scores in Texas tell us? *Education Policy Analysis Archives*, 8(49), Available: <http://epaa.asu.edu/epaa/v8n49/>
- Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.
- Kopriva, R. (2000). *Ensuring accuracy in testing for English language learners*. Washington, DC: State Collaborative on Assessment and Student Standards LEP Consortium, Council of Chief State School Officers.
- Koretz, D. M., Linn, R. L., Dunbar, S. B., & Shepard, L. A. (1991, April 5). The effects of high-stakes testing on achievement: Preliminary findings about generalizations across tests. Paper presented at the American Educational Research Association, Chicago.
- Lapp, D., & Flood, J. (2000, Fall). An interview with Marion Joseph. *The California Reader*, 34(1).
- McCarrier, A., Pinnell, G. S., & Fountas, I. C. (1999). *Interactive writing: How language & literacy come together, K-2*. Portsmouth, NH: Heinemann.
- McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass Publishers.
- Mora, J. K. (2001). What do the SAT-9 scores for language minority students really mean? Retrieved from the World Wide Web, January 26, 2001, <http://coe.sdsu.edu/people/jmora/SAT9analysis.htm>.
- Moustafa, M., & Land, R. (2001). The research base of Open Court and its translation into instructional policy in California. Retrieved from the World Web September 29, 2001, <http://instructional1.calstatela.edu/mmousta/>
- No Child Left Behind Act. (2001). Retrieved from the World Wide Web February 10, 2002, <http://www.ed.gov/nclb>.
- Office of Civil Rights. (2000). *The use of tests as part of high-stakes decision-making for students: A resource guide for educators and policy-makers*. Washington, DC: U.S. Department of Education.
- Ohanian, S. (1999). *One size fits few: The folly of educational standards*. Portsmouth, NH: Heinemann.

- Pyle, A. (1999, February 18). California and the West; School accountability bill clears panel; Legislature: Measure would allow state to rank campuses by performance and identify 300 for improvements. *Los Angeles Times*, pp. 3.
- Sacks, P. (1999). *Standardized minds: The high price of America's testing culture and what we can do change it*. Cambridge, MA: Perseus Books.
- Schneider, S., & Holtzman, D. (1999, November). *Evaluation of California's Standards Based Accountability System: Final Report*. San Francisco, CA: WestEd.
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1989). *The role of testing in elementary schools*. Los Angeles, CA: Center for Research on Educational Standards and Student Tests, University of California, Los Angeles.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Steck-Vaughn Company. (1998). *Test best for test prep*. Orlando, FL: Author.
- Stoskopf, A. (2000). *The forgotten history of eugenics*. In K. Swope & B. Miner (Eds.), *Failing our kids: Why the testing craze won't fix our schools* (pp. 76-80). Milwaukee, WI: Rethinking Schools, LTD.
- Swope, K., & Miner, B. (Eds.). (2000). *Failing our kids: Why the testing craze won't fix our schools*. Milwaukee, WI: Rethinking Schools, LTD.
- Taylor, D. (1998). *Beginning to read and the spin doctors of science: The political campaign to change America's mind about how children learn to read*. Urbana, IL: National Council of Teacher of English.
- Thompson, M., DiCerbo, K., Mahoney, K., & MacSwan, J. (2002). Exit to California? A validity critique of language program evaluations and analysis of English learner test scores. *Education Policy Analysis Archives*, 10(7), Retrieved from the World Wide Web January 26, 2002 from <http://epaa.asu.edu/epaa/v10n7/>
- Valencia, R. R. & Bernal, E. M. (Eds.). (2000). The Texas Assessment of Academic Skills (TAAS) case: Perspectives of plaintiff's experts [Special issue]. *Hispanic Journal of Behavioral Sciences*, 22(4).
- Wells, F. (2001). ZIP codes shouldn't determine our students' future. *California Educator*, 5(8), 1-5.
- Yanow, D. (2000). *Conducting interpretive policy analysis*. Thousand Oaks, CA: Sage Publications.

2002 Article Citation

Wright, W. E. (2002, June 5). The effects of high stakes testing in an inner-city elementary school: The curriculum, the teachers, and the English language learners. *Current Issues in Education* [On-line], 5(5). Available: <http://cie.ed.asu.edu/volume5/number5/>

Author Notes

Wayne E. Wright
Arizona State University
wayne.wright@asu.edu.

Wayne E. Wright is a doctoral student and graduate research assistant in the Educational Leadership and Policy Studies Program at Arizona State University. He currently serves as the Assistant to the Director of the Language Policy Research Unit of the Education Policy Studies Laboratory. He is also an editorial assistant with the *Journal of Language, Identity, and Education*, and recently served as an editorial assistant with the *Bilingual Research Journal*. Prior to coming to Arizona State University, Wayne was an elementary school teacher (grades K - 2nd) in Long Beach, California. Wayne taught in one of the only true Cambodian (Khmer) bilingual education programs in the country. He also served as a mentor teacher, and was the head of his school's technology committee. He lived and worked for over a year in Cambodia where he taught at the Institute of Economics, and provided technical assistance and training for human rights and student organizations. Wayne's research interests include bilingual education, standards and testing, and policy issues affecting the education of language minority students.

Note from the 2015 Executive Editor, Constantin Schreiber

May 26, 2015. This article was first published at the original *Current Issues in Education* website, located at <http://cie.asu.edu/articles/index.html>. In 2009, *CIE* changed online platforms to deliver the journal at <http://cie.asu.edu>. The original *CIE* website was from then on only used as an archival repository for published articles prior to Volume 12. After the new *CIE* website moved to a different server in 2014, the original website and original article URLs could not be accessed anymore. Therefore, this article had to be repurposed into the published format you are viewing now.

All content from the original publication has been preserved. No content edits occurred. Spelling, grammar, and mechanical errors that may be found were present in the original publication. The *CIE* logo and publisher information in use at the time of the article's original publication is unaltered. Please direct questions about this article's repurposing to cie@asu.edu.

2015 Article Citation

Wright, W. E. (2002). The effects of high stakes testing in an inner-city elementary school: The curriculum, the teachers, and the English language learners. *Current Issues in Education*, 5(5). Retrieved from <http://cie.asu.edu/ojs/index.php/cieatasu/article/view/1622>



Current Issues in Education

Mary Lou Fulton College of Education
Arizona State University