What Legitimate Inferences Can Be Made From The 1999 Release Of SAT-9 Scores With Respect To The Impact Of California's Proposition 227 On The Performance Of LEP Students?

Kenji Hakuta, Yuko Goto Butler and Michele Bousquet[1]

Stanford University

The recent release of California's SAT-9 test scores for the 1998-1999 academic year has attracted a great deal of attention, as these scores included the results of the first cohort of LEP students to be tested since the June 1998 passage of Proposition 227. In an effort to minimize confusion and to place some perspective on the situation, we have put forth below our interpretation of what can legitimately be gleaned from the data and, most importantly, what this means for the future educational practices and policies relating to LEP students.

Exclusion from the data of high-scoring upper grade non-native English speakers

SAT-9 scores for LEP students did increase somewhat from 1998 to 1999, especially in Grades 2 and 3 across the board. This increase can be seen in statewide scores for LEP students. For example, a 4 percentile point increase in reading and a 7 percentile point increase in math were made by LEP 2nd and 3rd grade students from 1998 to 1999. Similar increases can be seen as well as in the scores for all students, as depicted in Tables 1 and 2 below. (See http://www.stanford.edu/~hakuta/SAT9 for a more complete breakdown of all scores discussed in this article.) It should not alarm anyone that the scores for LEP students did not increase as much in the higher grades; as students get into the higher grades, those who do well on SAT-9 (and other measures of English proficiency) are "redesignated" into non-LEP status. Therefore, scores for these students are not included in the LEP data, causing students who perform well on these tests to be weeded out of the statistics.

Table 1. Statewide LEP Students' Reading Percentile Scores

Grade	1998	1999	Change
2	19	23	+4
3	14	18	+4

[1] The following students also contributed to the data analysis: Evelyn Orr, Jacob Mishook, Susan Baker, and Elsa Schirling.

Table 2. Statewide All Students' Reading Percentile Scores

Grade	1998	1999	Change
2	39	43	+4
3	36	40	+4

Score increases for districts maintaining bilingual education

Importantly, while rises can be seen in scores for LEP students in districts that claim to have faithfully implemented Proposition 227, such as Oceanside, rises can be seen as well in districts that have maintained various forms of bilingual education, such as Vista Unified, Santa Ana Unified, and Ocean View Elementary School Districts. At present, there is no scientifically defensible way to compare districts that have implemented Proposition 227 and those that have maintained bilingual programs. Both show positive changes, especially in 2nd grade. Table 3 below depicts 2nd grade LEP students' percentile point increases in reading from 1998 to 1999 in these school districts.

Table 3. Comparison of 2nd Grade LEP Students' Percentile Scores in Reading for Oceanside School District and Selected Districts Maintaining Bilingual Education

District	1998	1999	Change
Oceanside City Unified	12	23	+11
Vista Unified	18	26	+8
Santa Ana Unified	17	23	+6
Ocean View Unified	17	27	+10

Score increases for districts that never had bilingual programs

Rises in SAT-9 scores can also be seen for LEP students who have been in districts that have never had bilingual education programs. These districts had English-only programs in the year that Proposition 227 was passed and were therefore not impacted by Proposition 227's virtual elimination of bilingual programs. These schools with exclusively English-only programs include Orange Unified, Magnolia Elementary School District, Westminster Elementary School District, and Evergreen Elementary School District. Percentile point increases in reading for 2nd and 3rd grade LEP students in these schools can be found in Table 4. Although each of these

districts experienced a rise in scores from 1998 to 1999, these increases cannot be attributed to Proposition 227 since language of instruction was not changed in the schools within these districts.

Table 4. 2nd and 3rd Grade LEP Students' Percentile Scores in Reading for Selected Districts That Have Never Had Bilingual Education

District	Grade	1998	1999	Change
Orange Unified	2	16	23	+7
	3	15	16	+1
Magnolia Elementary	2	18	21	+3
	3	12	16	+4
Westminster Elementary	2	25	33	+8
	3	17	20	+3
Evergreen Elementary	2	50	54	+4
	3	30	42	+12

Score increases for low scoring native English speakers

Dramatic rises in SAT-9 scores can be found for native speakers of English at these grade levels in schools that registered very low SAT-9 scores in 1998. Focusing on 3rd grade data, our research group randomly identified 30 schools using State Department of Education data from 1998 in which there were fewer than 3% LEP students, but in which the average National Percentile Rank score was low (<27th percentile) in Reading for 1998. We tracked the changes in these schools for 1999. The data show an average increase of 8 percentile scores from 1998 to 1999 in these schools in reading. Similar gains were also seen in math and language scores for these schools. (See our website for a Table containing all 30 schools' scores. At least some of the gains seen in these schools is probably attributable to a statistical phenomenon known as “regression to the mean” in which scores at the extreme ends of the statistical distribution move toward the population average (mean), such that low scores move higher, and high scores move lower. The gains experienced in these schools are certainly not attributable to Proposition 227, since these are native English speakers in schools where almost all of the students are not limited-English-proficient.

Proposition 227 proponents' "model district" has below average score gains

Rises can also be seen in SAT-9 scores for LEP students at 3rd grade. We randomly sampled 26 schools that had high proportions of LEP students (greater than 80%), and who had low

reading scores in SAT-9 for 1998 (less than the 10th percentile). We also traced the changes in these schools for 1999. The results show an average increase of 4 percentile scores from 1998 to 1999 in these schools in reading, and comparable gains in math and language. (Refer to the website for a table listing the scores of each of these 26 schools.) We do not know the extent to which these schools implemented Proposition 227. However, it is notable that Oceanside Unified School District, which has been proposed as the model for Proposition 227 implementation, showed a gain of 3 percentile points in reading for 3rd grade (from the 9th to the 12th percentile). This is below the average gain found in this sample of schools.

Incorrect reporting of the data by Proposition 227 proponents

A final cautionary note is in order regarding Proposition 227 proponents' interpretations of its effect on the SAT-9 percentile point increases this year for LEP students. Proponents of Proposition 227 have incorrectly claimed on their website (see http://www.onenation.org) that "the Oceanside test scores revealed ... average percentile increases ranged from 120% in mathematics to over 180% in reading." We have been unable to determine exactly how they arrived at these figures. Taking even the most optimistic picture to be found in the Oceanside data, the very highest percentile increase in math for the 2nd grade is from 18 to 32 (a 14 percentile point increase) and from 12 to 23 (an 11 percentile point increase) in reading for the 2nd grade. None of these increases even in these best-case scenarios approaches the claim about a 120% to 180% increase.

The proponents' claim is probably based on making a comparison of the 1999 percentile score with the 1998 percentile score (i.e., for the increase from 12 to 23, one might divide 23 by 12, and come up with about 190%). This method, however, is incorrect. Without delving too deeply into the realm of statistics, it should be noted that when one starts with a low base, any increase will end up as a much higher percent increase. For example, a school starting at the 50th percentile (the national average) that goes up the same amount of 12 percentile points to 62, using the same division, will show only a 124% increase. By the same token, for a school at the 50th percentile to have the same amount of increase of 190%, it would have to increase its score to the 95th percentile! And, to carry it to the extreme, a school going from a percentile score of 1 to 2 (not a very respectable level of achievement) would have a 200% increase. In summary, if one starts low, then one does not have to go up very much to show a high rate of increase. As this relates to our discussion of SAT-9 scores, if increases are reported using a ratio of scores, any increase they make will seem larger than those made by non-LEP students that generally had higher initial scores. Does this mean that LEP students increased more than native English students, and therefore that we should accept the claim of a resounding success for Proposition 227? Of course not. Referring back to Tables 1 and 2, we see that both groups have increased by an *equivalent* 4 percentile points, the correct way to relay data of this nature. Even more refined would be to take a comparable sample of native English speakers who had low scores, as we did in our analysis above – in this case, native English speakers in reading in 3rd grade increased an average of 8 percentile points, compared to an average of 4 percentile points for LEP students. There is no conclusion to reach other than that supporters for Proposition 227 will have to look elsewhere if they want to advocate their case.

Conclusions

The conclusion we reach from this pattern is as follows: the increases in LEP students' scores for SAT-9 from 1998 to 1999 need to be considered in light of the overall gains in scores found across the state for all students. LEP students rose, as did non-LEP students. LEP students in English-only programs rose, as did LEP students in bilingual programs. And, native English speakers in low-performing schools made gains, as did LEP students in low-performing schools. These gains are probably the result of a combination of things. The fact that schools and districts have gotten used to the tests and are taking them more seriously should be considered (this is typically found in the second year of testing programs, as is the case for SAT-9 in California -- last year was the first time), as well as the fact that a variety of other initiatives such as class-size reduction may be taking effect. Additionally, it should be noted in the case of the low-scoring schools, that statistically, low scores tend to rise because of a "regression to the mean". Finally, when analyzing these scores, it is important to note that there are also a host of other uncontrolled factors that are involved in score fluctuations.

The policy conclusion we reach is that no one should be delighted by the fact that the overall performance of LEP students and of poor, native English speakers is very low on these standardized tests. These data should be mined further to determine why increases and decreases happened, and we should learn from the instances where high achievement can be found. We are delighted, however, that policy makers and the public, because of these data, have become concerned about the achievement of LEP students, and we are hopeful that this will lead to a deep and profound inquiry into how we can do better for the students. Better measures of English proficiency, such as those being proposed in California based on the English Language Development Standards, are also called for. We have long argued that focusing exclusively on whether one should teach only in English or using the native language is a major distraction that occurs at the expense of coming to serious grips with how to improve schools (August & Hakuta, 1997). We hope that this experience with trying to interpret the most recent release of SAT-9 data will convince the public that we should stop pointing the finger at bilingual programs and get into a serious discussion of improving schools, whether they be English-only or bilingual.

References

August, D. & Hakuta, K. (1997). Improving Schooling for Language Minority Students: A Research Agenda. Washington, DC: National Academy Press.