

High Schools and High Stakes Testing in California: Size and Income Do Matter

By:
L. D. Rector

Publication date
2/5/11

Abstract

The purpose of this study was to examine the relationship between the size of high schools, their percentage of SED (socio-economic disadvantaged) students, and API (academic performance index) scores in California, and determine if teacher preparation is a contributing factor. The 2010 API scores and median income of all 52 counties, and the 2010 API scores and %SED of 1089 high schools were tabulated and graphed to determine the strength of the correlation between the two different sets of data. Also, the percent proficient levels (in English) for all high school students by grade (9-11) and by socio-economic status from 2003 to 2010 were compared. Lastly, the number and percent of English and math teachers with the proper credentials are presented for analysis. Results indicate there is a strong correlation between the %SED and API of large high schools, and the correlation decreases as the size of high schools decreases. Also, the percent proficient levels in English of SED students are consistently lower than non-SED students across all grades. Results show that the performance of large, high %SED high schools on the CST (California Standards Test) is inferior to their large, low %SED counterparts, regardless of the percentage of teachers holding the proper credentials. Parents, policymakers, educators, and researchers need to question the scope and purpose of high stakes tests if the tests, by design, cannot account for circumstances beyond the control of students, parents, and schools. (Contains 5 graphs)

I'm tired of witnessing and contributing to my profession's consistent and persistent failure to tell the public the entire story. After more than 10 years as high school teacher, vice principal, and principal, I want the public to have in the form of verifiable data what we practitioners know anecdotally and many people know intuitively. In our profession, people always want to know "what the data says", often discrediting what many know from decades of direct experience and instead favoring the direction of someone with peripheral experience. The latter can come in the form of politicians, "experts", academics, writers, reporters, etc. To the dismay of many (or most) of us practitioners, and it's a reality that we subjugate ourselves to, these "latter" folks are the ones that drive educational policy and essentially define our jobs, profession, and consequently defining the treatment of your children. Ours is a profession where practice loses out to theory, and this study is but a humble attempt at consolidating both; there actually is data to validate what we know to be true.

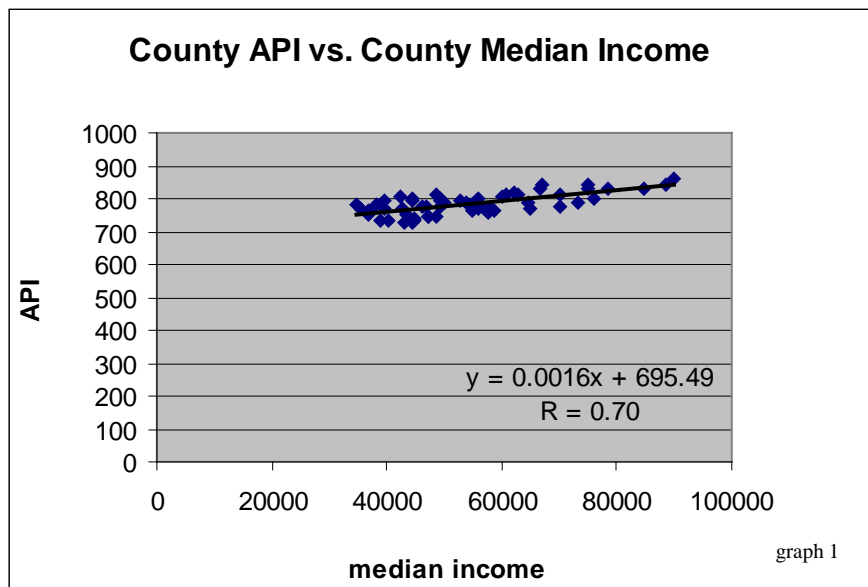
Many of us hate state testing in its current manifestation; that will be the first lie I'll disprove in case people think we are as hyped up about it as the sound bites make us out to be. What we hate about state testing is not the test itself, as educators we realize that assessment is an integral part of evaluation; we hate what state or "high stakes" testing has done to our profession; we're one big test prep. You've heard the complaints before, but mainly testing has sent the message that schools, and ultimately teachers, cannot be trusted to designate our students as proficient. It's not enough for students to complete their coursework and get their diploma; schools can no longer be considered "good" without the number to justify it. In California, this number is the API (Academic Performance Index), and as the data below will show, the number is no less biased and/or subjective than the teachers it was designed to objectivize. Our large, high poverty high schools and their students are, as you will see, being subjected to an institutionally imposed stigma that is masked in the name of student achievement and teacher accountability.

The state spends millions upon millions of dollars on high stakes testing, getting it right, and making sure its use is justified; you figure it should since many people lose jobs on account of it and many schools, teachers, and students have to bare the brunt of the public humiliation if they don't do well. A detailed description of the process followed in test development, administration, and reporting is beyond the scope of this paper, but you can read the **California Standards Tests Technical Report Spring 2009 Administration** for a detailed 620 page explanation of the test. One of the explanations you will find in it is the process of item (test question) review, which comprises of three levels: internal content review, internal editorial review, internal sensitivity review (pg. 199). The first note of interest for this paper is the following, "These assessment specialists make sure that the test items and related materials are in compliance...for clarity, style, accuracy, and appropriateness for California students..." (pg.199). The next note relevant to this paper is, "In their examination of test items, the ARPs [Assessment Review Panels] may raise concerns related to...socioeconomic bias" (pg. 200). Please keep in mind these two points as you look at the data that follows, particularly the second point.

With so much money, time, and expert care that are invested in these high stakes tests, then clearly the information presented about the API in the media must be accurate

and credible, right? My hope is that you draw your own conclusions after you review the data; it's time to stop relying on "experts" to tell you what is going on with testing and schools. Too many of us surrender our philosophical positions, and quite frankly our common sense, about the education of our students because someone that supposedly knows more has declared that we are wrong. The following information reflects the 2010 API scores and 2008 median income of all 52 counties, and the 2010 API and percent of socio-economic disadvantaged students of 1089 high schools in the state. The information will be presented to you in the form of graphs and explanations for you to make an informed decision and take a defensible position; not for someone to do it for you. This was not a study done with surveys, interviews, or observations in schools where many factors can influence the end result. All of the data is public knowledge and easily attainable as it was pulled from the California Department of Education (CDE) website (www.cde.ca.gov) and the county median income levels from the US Census Bureau website (www.census.us.gov); anyone with the inclination and time to replicate these graphs can easily do so.

Please remember that the high stakes CSTs (California Standards Test) are extensively and expensively reviewed so as to "...raise concerns related to socioeconomic bias." Because of this, we should expect the test to favor no student or school on account of socioeconomic level. Our first evidence that the test fails in this respect is in the review of county wide 2010 API scores as illustrated in graph 1 below.

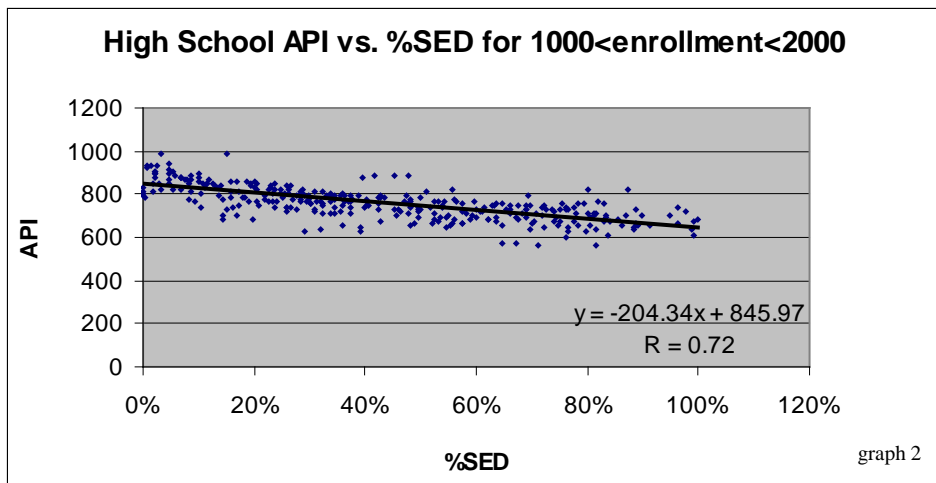


This graph shows the relationship between the average API and median income levels of each California county. The average API was calculated by taking the API of every district in the county and averaging those district APIs. The only scores not included in the county averages were those of county offices of education and any districts without valid scores reported in 2010. The median income levels were from 2008 as those were the last published by the US Census.

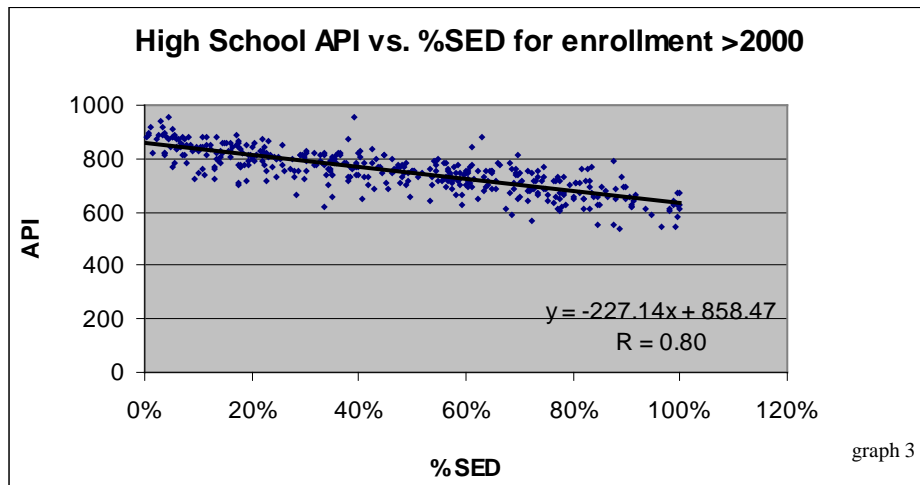
Clearly, there is a relationship between the median income and API; the API increases with increasing median income. For those with a mathematical inclination, I included the correlation coefficient; the "R" in the lower right of the graph. Put simply, an R value close to 1 means that the line is a good predictor, and .70 is a fairly strong

correlation. What is impressive about the .70 value is that counties can be large entities with pockets of high and low income areas dispersed throughout and yet we get a strong correlation between the income level and the API scores; one would expect the data points to be scattered throughout the graph.

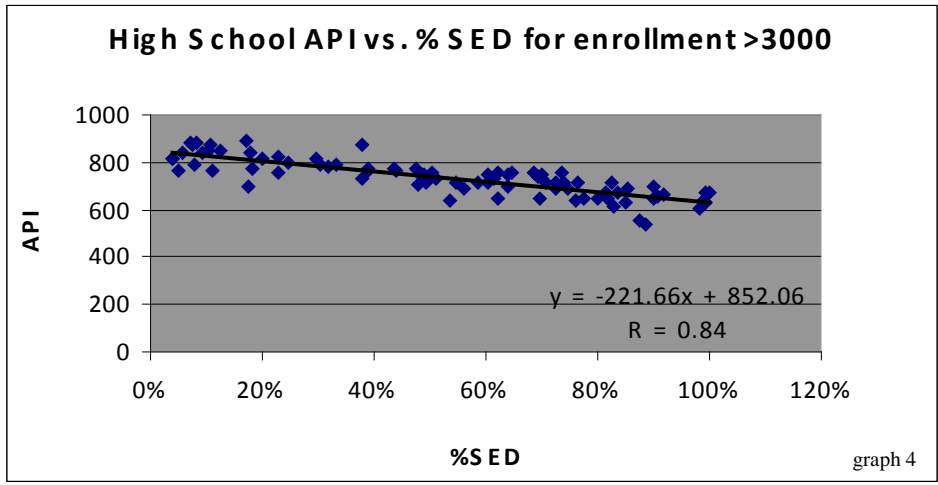
Here we have our first indicator that the CSTs may not have been designed to overcome socioeconomic bias for certain types of schools. If something as large and economically heterogeneous as a county can give such a strong correlation between API and income, then how does it play out for individual high schools? The following graph (graph 2) is of 321 high schools with populations of 1000 to 2000 students. As the CDE does not collect and publish information on median income, the income information used to measure against the API is %SED (Socio Economic Disadvantaged). The CDE defines a student as SED as one that qualifies for the Federal Free and Reduced Lunch Program. The %SED is the percent of students in a school that qualify for free and reduced lunch.



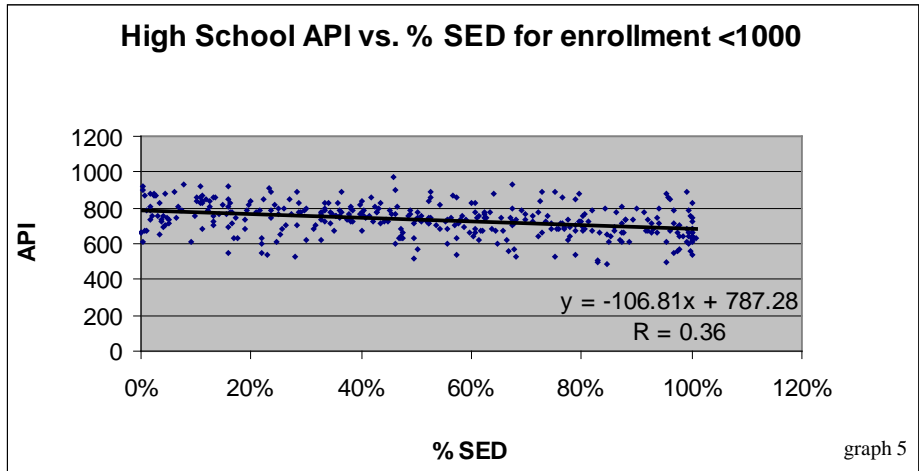
The answer to our question at the beginning of the previous paragraph is obvious; there is still a strong correlation (stronger than the county data) between the socioeconomic status of students and schools and the API. Here, however, the relationship is negative; the greater the %SED, the lower the API. The following graph (graph 3) is of 410 high schools with populations of more than 2000 students.



There is a great jump in the correlation between API and %SED. The following graph (graph 4) is of 83 high schools with populations of 3000 or more students.



Yet again we see a jump in the correlation between API and %SED; the predictability of the performance of a high schools increases tremendously the larger it gets. Something interesting, though, happens with high schools with less than 1000 students. The following graph (graph 5) shows the relationship between API and %SED for 358 small high schools.



What does this data tell us about the relationship between API and %SED for large high schools? Clearly, the larger the high school, the greater likelihood that a large %SED population will result in a lower API scores. If the CSTs “are in compliance...for... accuracy, and appropriateness for California students...” and specialists are used to “raise concerns regarding socioeconomic bias”, then why is there a clear effect of socioeconomic status on the API? The results should look like those of graph 5; there should be no predictable relationship between the %SED and API if the test is truly to be considered fair and unbiased. A school should be able to attempt the test without knowing with a high degree of certainty that they may not score well. Those of us that work in large, high %SED schools know and experience, everyday, the effects of poverty on the students and schools, but we are reminded by “experts”, politicians,

researchers, etc., that poverty or school size is no excuse for the failure of our students and our schools. Everyday we accept the challenge; however, if the same instrument used to measure our success, CSTs, has an inherent bias then how fair is it to use that faulty instrument? If the CDE spends millions of dollars in creating and administering this high stakes test and holds schools accountable to the point of creating stigmas and eliminating people from their positions, then a minimum expectation from all of us should be that the test is free of bias so that it shouldn't matter if a school has many of few low socioeconomic students. Clearly, the test fails in this respect if we hold the test to the same standard that we hold our schools and teachers to, that poverty is no excuse. So, if it's not the test, and it's not poverty, then what is it?

Traditionally, and conveniently, the finger has always been pointed at the schools, and specifically at teachers. However, if we look at the teacher data objectively and use a little reasoning, one finds that teachers may be carrying the brunt of the blame unjustifiably. The latest teacher certification information that the CDE provides is from 2009, and it shows that 93.6% of high school English teachers in the entire state hold the correct credentials, highly qualified is the federal designation, to teach high school English, and 91.76% to teach high school math. Let's consider the large high schools (just to start) with enrollments greater than 2000 and an API lower than 700, which is considered quite low and subject to sanctions. From 1089 high schools in the data set, we find that there are 101 of those schools. According to 2009 data from the CDE, there were 2493 English 9 teachers, 1910 English 10 teachers, and 1309 English 11 teachers (the three grades that test) for a total of 5712 English teachers. If 93.6% of teachers held the correct credentials (highly qualified), then that means that 6.4% presumably were not highly qualified. A simple calculation gives us that 6.4% of 5712 is just over 365. In schools larger than 2000, English departments have about 20 teachers (obviously more for the much larger schools). So, our 101 schools with APIs less than 700 collectively account for, at least, 2000 English teachers, far more than the 365 "not highly qualified". Is it reasonable then, to conclude that in ALL of these 101 low performing schools the teachers are ineffective even though statistically the vast majority is highly qualified by federal standards? Even if we distributed evenly all of these "unqualified" teachers so that some schools had departments full of them, then that would only amount to about 20 of these large schools; even the most reticent skeptic knows that will not happen. Let's go the other way; is it reasonable to conclude that none of the 310 large high schools with scores greater than 700 have ANY of these "unqualified" English teachers?

If we take the same line of reasoning for math, without having to go into great detail, we will come to the same conclusions. I will provide the data for Algebra 1 only, as it is the only math course that the state requires of everyone for graduation and math classes in high school do not fall into delineated categories (like English 9, 10, 11); you can have some 9th graders taking Algebra 2 while some 11th graders are still taking Algebra 1, for example. In 2009, there were 5071 Algebra 1 teachers with 91.76% of all secondary math teachers classified as highly qualified as described above; do the math like we did above and you will arrive at a similar conclusion.

There is one last bit of proof that maybe, just maybe, the source of the problem could be the test. We have spent the entire paper up to this point analyzing the negative relationship between the size and %SED of high schools and API. There are many large high schools that have already met the 800 API target. Although the data shows that these

schools tend to have a lower %SED, that does not take away any well deserved merits; just running a large schools in and of itself is a daunting task. However, the high performing schools and the parents of the students that attend them should not succumb to the temptation of blindly refuting the argument that there may be something wrong with the CSTs; they should “not get too happy”. In a study conducted by the American Institutes for Research, the authors sought to answer the question, “How would the 2007 state results reported to No Child Left Behind have looked had all the states used a common performance standard that had been internationally benchmarked to TIMSS or PIRLS?” (pg.15). In other words, the researchers set out to determine how our state standards across the nation would compare to the standards of countries that have traditionally outscored the United States. For California 8th grade math, the percent proficient would only be 22; by comparison, Massachusetts would have a percent proficient of 52. Keep this in mind as you read the very first paragraph of the CDE technical report cited earlier:

*In 1997 and 1998, the California State Board of Education (SBE) adopted **rigorous content standards** in four major content areas: English–language arts (ELA), mathematics, history–social science, and science. These standards are designed to guide instruction and learning for all students in the state **and to bring California students to world-class levels of achievement.*** (emphasis added)

The authors of the study concluded that “for Grade 8 mathematics, Massachusetts and South Carolina were the only states with world-class standards.” (pg. 15)

Summary

As promised earlier, I leave you with the data so that you can draw your own conclusions. The reflection questions presented throughout the paper were provided to assist you in moving from one way of understanding something to another; in teaching this is called scaffolding. I leave you with more questions to ponder and provide more scaffolding.

- High stakes tests are designed with a high degree of reliability; high degrees of reliability help ensure that if students were to take the same test different times then their scores will not change (increase or decrease) a great deal. With such a high degree of reliability and correlation between size and %SED, what can we expect of the scores of large, high %SED high schools?
- If we cannot predict the scores of small, high %SED high schools, then couldn't we look at what other factors are contributing the low scores?
- If we cannot predict the scores of small, high %SED high schools, then why don't we have more?
- If we can predict the scores of large high schools by looking at the %SED, why do we continue to compare the ones with low %SED with the high %SED ones?
- If we can predict the scores of large high schools by looking at the %SED but not the small ones, then should the intended use of the CSTs be changed?
- If we can predict the scores of large, high %SED high schools, and knowing the sanctions and stigmas associated with low scores, then why would teachers want to work there?

- If teachers are so bad at those schools and the ones at the large, high performing ones are so good, and teachers are to blame for the low performance, then why would it take an act of congress and relentless negotiations with teacher unions to swap them?
- If teachers are so good at the large, high performing high schools and so bad at the low performing ones, why doesn't someone offer an incentive to the good ones to work at a large, high %SED high school for the purpose of conducting a study to determine if that will make a difference?
- If teachers are so good at the large, high performing high schools, then why would they need an incentive to do that?
- If private (denominational and not) schools do not take the CSTs and there is no number to indicate their effectiveness, then why do people choose to pay money to send their children there?
- If private schools do not require teacher credentials, then why do people choose to pay money to send their children there?
- If something is free but you are forced to consume it, then is it really free?
- If we can predict the scores of large, high %SED high schools, then why don't we remind parents (the way we remind them about the scores) that they have a right to have their students opt out of the CSTs?
- If parents have a right to have their students opt out of the CSTs, then why will schools be sanctioned if too many opt out?
- What message are we sending by giving parents a legal right to have their children opt out of the CSTs but at the same time sanctioning schools if too many opt out?
- Do we really care about kids, all kids?