

Proposition 227 in California:
A Long-Term Appraisal of Its Impact on Language Minority Student Achievement

Laura McCloskey ¹	lauramc@stanford.edu
Nathan Pellegrin ²	nathan.pellegrin@gmail.com
Karen Thompson ¹	karendthompson@stanford.edu
Kenji Hakuta ¹	hakuta@stanford.edu

Fall 2008

Preparation of this paper was supported in part by a grant from the Hewlett Foundation to Kenji Hakuta.

We would like to gratefully acknowledge the feedback we received on earlier drafts of this paper from Joe Robinson, Patricia Gándara, Robert Linquanti, and Russ Rumberger.

¹ Stanford University

² Nathan Pellegrin and Associates

Abstract

This study examines the impact of Proposition 227 on educational outcomes for California's 1.5 million English learners. Building on prior research (Parrish, Pérez, Merickel, & Linqanti, 2006), we compare English learner and English-only student achievement by grade level, using data from the California Standards Test from 2003 to 2007. While both ELs and EOs show a positive trend in CST scores over time, there were differences between the slopes for the two groups in Grades 6, 7, and 8. Specifically, the 8th grade test score trend line for English learners shows a less positive slope than the test score trend line for English-only students, suggesting evidence of negative impact for ELs relative to EOs. However, the 6th grade test score trend line for ELs shows a more positive slope than the test score trend line for EOs, suggesting the ELs showed more relative progress over time than EOs. This pattern suggests that Prop 227 had a localized negative impact that was especially observable in the 8th grade data. However, in the subsequent 2 years of implementation, as detected in the 6th and even in the 7th grade data, the EL achievement seems to have increased relative to EO achievement. Further analysis comparing school districts in which ELs were performing especially better or worse over time relative to EO students found no statistically significant differences in instructional programs for English learners or in demographics across these two types of districts. However, limitations of the state data system, specifically the lack of student-level data about instructional services, as well as the lack of longitudinal data at the student level, limit the explanatory power of this analysis.

Introduction

In June 1998, California voters passed Proposition 227 (hereafter Prop 227), a law requiring public schools to conduct instruction “overwhelmingly” in English. With over 1.5 million students in California classified as English learners, the law affected one-fourth of California students and over one-third of English learners in the United States (California Department of Education, 2008a; U.S. Department of Education - NCES, 2006). The law especially impacted English learners receiving instruction in bilingual programs. Since the passage of Prop 227, the percentage of English learners receiving primary language instruction has decreased from 29.1% in 1997-1998 to 5.6% in 2006-2007 (California State Department of Education, 2008b; California State Department of Education, 2008c).

Prop 227 drastically changed California’s policies for educating English learners. Before Prop 227 passed, districts had control over the instructional methods they used to educate their English learners. However, under Prop 227, English-only instruction was mandated for all English learners unless parents signed a waiver otherwise. Prop 227 stipulated that English learners be placed in Structured English Immersion (SEI) classes in which instruction was required to be “overwhelmingly in English.” Once students acquired a “good working knowledge of English,” students transferred to mainstream classrooms.

For almost ten years now, school districts and more importantly English learners have felt the impact of Prop 227’s policy change. A number of research reports have attempted to analyze the impact of Prop 227 with varying methods and findings (e.g. Parrish, Perez, Merickel, & Linquanti, 2006; Gándara, Maxwell-Jolly, Stritikus, Curry, Garcia, Asato, & Gutiérrez, 2000; Butler, Orr, Guitierrez, & Hakuta, 2000). In most cases, the reports relied on achievement data that straddled three different standardized tests, the Stanford-9 Achievement Test (SAT-9), the

California Achievement Tests, Sixth Edition Survey (CAT-6), and the California Standards Test (CST). Building on the findings from the state-mandated report conducted by Parrish et al. (2006), this study uses five years of CST data to examine Prop 227's impact on English learner achievement.

Review of Literature

Many studies have addressed the impact of Prop 227 on language minority students. Some studies showed the rapid decline in the number of students receiving instruction in their primary language (Gándara, et al., 2000; Parrish, et al., 2006). Other studies found wide variability in the implementation of the new policy, including variability in districts' use of waivers, in teachers' instructional responses, and in community attitudes towards Prop 227 (Gándara, 2000; Gándara, et al., 2000; Maxwell-Jolly, 2000). Studies investigating the relationship between Prop 227 and student achievement found a variety of results. For example, some studies found that the achievement gap between native English speakers and English learners still existed and that Prop 227 did not increase the rate of English learners' reclassification as Fluent English Proficient (Butler, et al., 2000; Gándara, 2000; Grissom, 2004).

Parrish et al.'s (2006) study constitutes the largest, most thorough, and most recent examination of Prop 227's impact. Commissioned by the California Department of Education, the study lasted five years. It covered a broad range of topics, including the impact of Prop 227 on English learner achievement, the overall implementation of Prop 227, promising practices for English learners, issues around the reclassification of English learners, and the support of English acquisition in the community.

The Parrish et al. (2006) study focused heavily on investigating possible changes in the achievement gap between English learners and non-English learners following Prop 227's implementation. Parrish et al. (2006) analyzed the test scores of students in all four language proficiency groups, as defined by the state: English learners (ELs); former English learners who have been redesignated as Fluent English Proficient (RFEPs); students who only speak English (EOs); and students who speak another language in addition to English but were initially classified as fluent in English when they entered school (IFEPs). To track achievement for these four language proficiency groups, Parrish et al. calculated achievement scores based on data from the Stanford Achievement Test, 9th Edition (SAT-9) from 1997-2002 and on data from the California Standards Tests (CST) from 2002-2004. The authors addressed the challenge of having different tests with different scales by using a metric-free measure to gauge gains over time.

From their analysis, Parrish et al. (2006) found a slight decrease in the performance gap between native English speakers and English learners in the years following Prop 227's implementation, but in general, the performance gap remained constant in most subject areas for most grades. When students who were former English learners were included in the English learner cohort, the performance gap remained. Also, the study looked at state and district data in the Los Angeles Unified School District to examine English learner performance under different instructional models, including bilingual and English-only programs. This analysis found little to no evidence of differences in performance for English learners by model of instruction. Our study builds on Parrish et al.'s (2006) methods and findings regarding Prop 227's impact on English learner achievement.

Study Design, Methods and Results

Prop 227 intended to improve academic outcomes for English learners through the elimination or reduction of bilingual education programs. Voters passed the initiative in June 1998, and the law took effect immediately, impacting instruction beginning in the fall of 1998. Figure 1 shows the years of Prop 227's impact, as well as the various standardized tests used in the state testing program over time. In this figure, we can see the successive cohorts of students who experienced these policies and the years in which the state administered different standardized tests. The dotted line in Figure 1 separates pre- and post-227 cohorts. All students above the dotted line entered school after Prop 227 was passed, whereas those below the dotted line had various amounts of schooling prior to Prop 227, depending on their grade levels when Prop 227 took effect.

[INSERT FIGURE 1 HERE]

Using available CST data from the five-year period from 2003 to 2007, we examined student achievement separately by grade level slices. For each district, we estimated the test score trend line separately for ELs and for EOs at each grade level. Comparing ELs to EOs within the same district serves as a control for district-level effects. Of particular interest was whether the EL and EO slopes diverged. If ELs benefited from Prop 227, we predicted that the test score trend line for ELs would have a more positive slope than the trend line for EOs, resulting in a positive slope difference. However, if Prop 227 harmed ELs, we predicted a negative slope difference, with the trend line for ELs having a less positive slope than the trend line for EOs.

This prediction would vary somewhat for each grade level slice in our analysis, since Prop 227 affected students in each grade level slice at different times in their educational

trajectories. Since most bilingual programs served students only in grades K-3, Prop 227 had the greatest impact on educational programs for students at these grade levels. The Grade 8 data slice included cohorts that were in Grades K, 1, 2, 3 and 4 in the year that Prop 227 was first implemented and therefore contains one cohort (from the 2007 test year) that entered kindergarten the year of Prop 227's implementation and never experienced a pre-Prop 227 educational environment. The test data from each descending grade level slice contains an additional cohort that was fully post-227. The cohort distribution and some predicted slopes are illustrated in Figure 2.

[INSERT FIGURE 2 HERE]

For the analysis, a dataset was created from the public state research data files, using districts as the unit of analysis. Our selection criteria required districts to have a mean scaled score for each combination of subgroup (EO/IFEP and EL/REFP), subject area (ELA and Math), test grade level (grades 2 through 8 for ELA, and grades 2 through 7 for math), and administration year (2003 through 2007), for a total of 130 data points for each district. Across the state research files, 994 school districts were found to have at least one data point. Of that number, 359 (36%) were found to have a complete set of data. The dataset analyzed in this study consisted of these 359 districts. The EL subgroup included data for only those English learners who had been in the system longer than a year in order to minimize the effects of students who were recent arrivals and therefore not affected by Prop 227. In the state dataset, performance of RFEPs and IFEPs are reported separately. The RFEP data were added to the EL data, since RFEP students had been classified as English learners at an earlier point in time. Data for the IFEP students (a small group) were added to the EO data, since this group of students, though language minorities, were proficient in English at initial assessment and never classified as ELs.

In both cases, the adjustments to the EL and EO numbers were weighted for the number of test takers in each group. The set of districts analyzed in this study represent 90% of all tests administered to EL/RFEP students, and 81% of all tests administered to EO/IFEP students statewide within the five-year period.³

Next, for each district, linear regression was used to obtain a slope of the CST score on year of testing, separately for the EO and EL groups. Table 1 and Figure 3 display the results from this slope analysis. The slopes for the EO and EL groups represent the average change in each group's mean CST score over time. There are two obvious and expected characteristics of the slopes that nevertheless need to be noted. First, the slopes are positive across all grade levels, and for both the EO and EL groups (see left panel of Table 1), indicating that CST scores for both groups rose over time. Second, there is a gap in the mean scores of the EO and EL groups on the order of 20-30 CST points. But our focus of interest is whether the slopes of the trend lines for EOs and ELs are different. The slopes of the trend lines for EOs and ELs represent the change in each group's mean CST scores over time. The differences in slopes can be seen in the extent to which the slopes are not parallel (see Figure 3). Additionally, these differences in EO and EL slopes are expressed numerically (see right panel of Table 1). Positive slope differences mean that the EL slope exceeded the EO slope, while a negative slope difference means that the EO slope exceeded the EL slope. Positive slope differences were found for Grades 3, 5, 6 and 7. Negative slope differences were found for Grades 2, 4 and 8.

³ In all subsequent mentions of language proficiency groups, the EL group refers to the combined group of English learners plus Redesignated Fluent English Proficient students. Similarly, future mentions of the EO group refer to the combined group of English only students plus Initially Fluent English Proficient students.

[INSERT FIGURE 3 HERE]

Before interpreting these patterns, a test of significance for the difference in slopes was performed, using a correlated samples t -test (since the EO and EL slopes are non-independent, i.e., correlated across districts, ranging from $r = .32$ to $.48$ depending on the grade level). These results are also displayed in Table 1. The significance test is simply performed against the null hypothesis that the mean of the differences in slope is zero.

[INSERT TABLE 1]

We are interested particularly in the slope differences for Grades 6 and 8, since they have the largest magnitude and are in opposite directions. For both grades, the effect size in standard deviation units was $.37$ (Grade 6 $M = -.900$, $SD = 2.456$, Grade 8 $M = .883$, $SD = 2.390$), suggesting a moderate-sized difference averaged across the school districts. The fact that the slope differences at Grades 6 and 8 go in opposite directions demands attention. These results could be interpreted to suggest support for the two contrasting hypotheses suggested by the graphics in Figure 2 – that Prop 227 could perhaps be associated with “harm” to ELs and with “benefit” to ELs.

The Grade 8 results are of particular interest since every cohort in this grade level was enrolled in school the year that Prop 227 took effect (see Figure 1). Therefore, we might expect to observe Prop 227’s greatest impact with this grade level slice. If one assumes that any potential harm caused by pre-227 policies is best removed earlier rather than later, then one would expect a positive difference in slope over time. On the other hand, if a longer period of “treatment” under pre-227 policy was better, and Prop 227 caused a harmful disruption to this treatment for ELs, then a flatter slope would result. The results for the 8th grade slice show a flatter slope for ELs than EOs and appear to be consistent with the second hypothesis that Prop

227 was harmful for ELs. Furthermore, this apparent harm is detectable downstream in students' educational careers, after considerable time has elapsed.

The Grade 6 results show the opposite pattern, with a greater increase in test scores over time for EL students relative to EO students. A key difference between the 6th grade and 8th grade slices is that the 6th grade slice includes two cohorts, from test years 2006 and 2007, who were not enrolled in school when Prop 227 took effect. (As noted above, all cohorts in the 8th grade slice were enrolled in school when Prop 227 took effect.) Referring to Figure 1 and looking at the Grade 6 testing years, it is evident that the 2003 and 2004 test results in the 6th grade data slice were for students who had experienced one to two years of schooling prior to Prop 227's implementation. One interpretation of the 6th grade results would be that the 2005, 2006 and 2007 cohorts were liberated from the old policy, and perhaps with better implementation of the new policy, they did better in successive years. One might argue that if bilingual education had been benefiting ELs, then the 2003 and 2004 6th grade ELs should have outperformed later cohorts of 6th grade ELs, relative to EOs, because they had benefited from two years or one year of pre-227 policy, respectively. In that case, a negative slope difference would be predicted. Instead, there is evidence that the 2003 and 2004 6th grade ELs performed more poorly, relative to EO 6th graders, than later cohorts did. On the other hand, advocates for bilingual education might argue that the one or two years of bilingual education that the 2003-2004 groups might have received is insufficient and perhaps could lead to poorer outcomes because these students suffered from inconsistent experiences in their education.

Examining Differences in Instructional Services for ELs Among Districts

To further shed light on the impact of Prop 227, we analyzed data at Grades 6 and 8 more

closely. Specifically, we identified school districts that had relatively large positive and large negative slope differences (i.e., those districts where the ELs were performing better or worse over time relative to EO students), and examined their reported instructional services for ELs both before and after Prop 227. The distribution of schools with respect to slope difference can be seen in the box plots in Figure 4, with the 359 individual districts in our dataset represented by small circles. As the box plots demonstrate, there is considerable variability in slope differences between groups at each grade level, with slope differences ranging from 13.516 to -7.794 at 6th grade and slope differences ranging from 7.679 to -8.589 at 8th grade.

[INSERT FIGURE 4]

From this distribution of districts, we selected the top 20 and bottom 20 districts, in terms of slope difference, at both 6th and 8th grades. We excluded districts that were outliers for EL slope and EO slope. We then retrieved data from the “English Learners, Instructional Settings and Services” section of California’s online Language Census database. This database documents the number of ELs in each district that were listed as receiving one of six categories of instructional services: English Language Development (ELD); ELD + Specially Designed Academic Instruction in English (SDAIE); ELD + SDAIE + Primary Language Support; ELD + Academic Subjects through the Primary Language; Other Services; and No Services.

We gathered data about the number of ELs in each category for each district at four points in time – 1997-98 (the year Prop 227 passed), 1998-99 (the year Prop 227 was implemented), 2002-03 (the first year the CST was administered), and 2006-07 (the most recent year for which data were available). While these data provide important information about instructional services provided to English learners district-wide at particular points in time, they have significant limitations. First, since we do not have student-level data, we do not know

which English learners in our 6th and 8th grade samples received particular instructional services. We also do not know the length of time for which particular English learners may have received particular services. Furthermore, since the instructional services data are not broken down by grade level, we cannot assess what percentage of a district's English learners received particular instructional services at specific grade levels. The instructional services data merely provide a snapshot of aggregate information about the number of English learners receiving particular instructional services and allow us to document whether there were large differences across districts in the percentages of ELs receiving different services over time.

Using the total number of ELs reported for each district at each point in time in the Language Census database, we converted the raw number of ELs receiving each type of instructional service into percentages of ELs receiving each type of service. Next, since percentages are bounded and not normally distributed, we conducted an arcsine transformation of these percentages, as outlined in Table 2. Then, using these transformed percentages, we calculated the means for English learners receiving different types of services at different points in time for the top 20 and bottom 20 districts at 6th and 8th grades. Next, we conducted *t*-tests to determine whether the differences between these means for the top 20 and bottom 20 districts were statistically significant.

Since Prop 227 attempted to improve educational outcomes for English learners by limiting their enrollment in bilingual programs, we were particularly interested in the percentages of ELs that were listed as "Receiving Academic Subjects through the Primary Language" for the top 20 and bottom 20 districts over time. Past research has used the percentage of English learners listed in this category as a proxy for the percentage of English learners enrolled in bilingual programs (e.g. Parrish et al., 2006). We will continue this practice.

[INSERT FIGURES 5 AND 6 AND TABLES 2 AND 3]

As evident in Figures 5 and 6, the most striking feature of districts' instructional services for ELs is the sharp decline in the percentage of English learners who received academic subjects through their primary language over time. This trend occurs at the state level and for both the top 20 and bottom 20 districts at 6th and 8th grades. While Parrish et al. (2006) documented this trend from 1997-98 through 2003-04, our data show that the percentage of ELs receiving academic subjects in their primary language has continued to drop, from a statewide average of 29.1% in 1997-98 before the passage of Prop 227 to a statewide average of 5.6% in 2006-07. While there are differences in the means of the top 20 and bottom 20 districts at both grade levels, these differences are not statistically significant, as demonstrated in Tables 2 and 3.⁴ Caution must be exercised when interpreting these results, given the limitations of the available data on instructional services. With that caution in mind, the analysis of the available data suggest that districts that differ in the test score trends of their EL students compared to their EO students do not differ significantly in terms of the instructional services they provide to English learners, as measured by the state data system. However, disaggregated, school-level or student-level data, would enable deeper analysis of differences in the instructional services English learners in our sample received.

Examining Differences in Descriptive Characteristics Among Districts

Since we found no significant differences between the top 20 and bottom 20 districts at 6th and 8th grades related to their instructional services for English learners, we turned to examining

⁴ We compared means for the percentage of ELs receiving other categories of instructional services, as well, and found no statistically significant differences between the top 20 and bottom 20 districts for these other services.

other descriptive characteristics of the districts. We collected information from the online state Enrollment database about the ethnicities of students in each district, as well as information from the Language Census database about the languages spoken by the English learners in each district. In addition, we used the Ed-Data website to collect information about additional variables for each district, including: total enrollment; percentage of students in the district that are classified as English learners; percentage of students who receive free or reduced lunch; percentage of teachers who are fully credentialed; percentage of schools in the district that made their Academic Performance Index goals under California's accountability system; percentage of students in the district that scored Proficient or above on the ELA section of the CST; percentage of students in the district that scored Proficient or above on the Math section of the CST; and whether the district met its Adequate Yearly Progress Goals, as defined by No Child Left Behind. Again, these data have significant limitations. We collected the descriptive data only for the 2006-07 school year, and districts' descriptive characteristics may have changed over time. Nonetheless, this descriptive data does provide us with a current snapshot of each district and could potentially illuminate current differences in characteristics of the top 20 and bottom 20 districts at 6th and 8th grades.

After gathering these data, we calculated means for each variable for the top 20 and bottom 20 districts, respectively, at both 6th and 8th grades. Figure 7 displays the means for the top 20 and bottom 20 districts at 6th grade for each of these variables, while Figure 8 displays the means for the top 20 and bottom 20 districts at 8th grade.

[INSERT FIGURES 7 AND 8]

To test whether the differences in the mean values for these variables were significantly different for the top 20 and bottom 20 districts, we conducted t -tests of the difference between the means,

after using the arcsine transformation if the variables were reported as percentages. We found no statistically significant ($p < .05$) differences between the top 20 and bottom 20 districts at 6th and 8th grades for any of the variables we analyzed. Again, caution must be exercised when interpreting these findings, given the limitations of this descriptive data. While the top 20 and bottom 20 districts at 6th and 8th grades do not appear to differ significantly in their descriptive characteristics for the 2006-07 school year, it is conceivable that significant differences in particular characteristics may have existed at other points in time.

Discussion and Policy Implications

Given our findings, has Prop 227 improved educational outcomes for English learners? While limitations in available data temper our ability to answer this question definitively, it is clear that on average, students in California, including English learners, have shown increases in standardized test scores since Prop 227's passage. However, many factors, besides any possible effects of Prop 227, are likely influencing this change in CST scores. For example, under No Child Left Behind and California's accountability system, schools and districts have federal and state CST performance targets and face sanctions if students' scores do not improve each year. Given the shifting, intertwined policy landscape, it is difficult to tie these general achievement trends to Prop 227. Furthermore, the significant achievement gap between English learners and English-only students, while perhaps closing slightly at some grade levels and widening slightly at others, remains substantial.

What could explain the conflicting CST trends at different grade levels? The Parrish et al. (2006) analysis of the effects of Prop 227 included this summary of its findings:

We conclude that Prop 227 focused on the wrong issue. It does not appear to be the model of instruction employed, or at least not the name given to it, but rather

other factors that are much more operative in distinguishing between failure and success with ELs (xii).

Our results support the conclusion of the Parrish et al. (2006) report. Available statistics about the instructional services offered for English learners by a district did not predict how trends for the districts' ELs' performance on the CST compared to the trends for its EOs.

What distinguishes districts in which the achievement gap between ELs and EOs narrowed from districts in which it widened? What might explain the differences in the comparative achievement trends of ELs and EOs at 6th and 8th grades? Why do we see an achievement gap between ELs and EOs that appears to be narrowing at 6th grade but widening at 8th grade?

Establishing definitive answers to these questions is impossible, but prior research may suggest possible interpretations. First, some studies suggest that it is not the particular type of instructional program that determines English learners' educational outcomes but rather the program's implementation (e.g. Gold, 2006; Parrish et al, 2006; Williams, Hakuta, Haertel, et al., 2007). While districts in our analysis had instructional services for ELs that appeared similar on paper, their programs for ELs may have differed on dimensions not captured by the state data system, such as expectations for students, instructional leadership, use of assessment, and linkages between home and school.

Second, the initial implementation of any major policy shift, particularly one as far-reaching and as abruptly implemented as Prop 227, likely causes upheaval, regardless of the merits of the program. The implementation of Prop 227 was further complicated by initial widespread uncertainty about the law's requirements. For example, approximately half of the 153 schools surveyed by Parrish et al. (2002) reported that they needed additional guidance from the state regarding what instructional arrangements were permissible under Prop 227.

The 8th grade slice in our sample was comprised of students who were in kindergarten through fourth grades when Prop 227 took effect. The confusion surrounding the law's implementation likely affected many students in this 8th grade slice. The students who were kindergarteners during the first year of implementation possibly entered schools in which there was confusion about waivers, about the amount of primary language instruction that was allowed, and about what "Structured English Immersion" meant. Meanwhile, the other students in the Grade 8 data slice, who entered Grades 1, 2, 3, or 4 during the year that Prop 227 took effect, may have been pulled out of bilingual programs that their districts eliminated.

The 6th grade slice in our sample, meanwhile, consisted of two cohorts of students not yet enrolled in school when Prop 227 took effect, plus three cohorts of students in kindergarten, first, or second grades that year. Given the findings of Parrish et al. (2002) regarding the challenges schools encountered in implementing Prop 227, it is likely that by the time the two later cohorts entered school, confusion over Prop 227 had subsided and districts were more likely to have the coherent, consistent programs in place – bilingual or not - necessary for English learners to experience academic success. Thus, the seemingly contradictory findings of a widening of the achievement gap between ELs and EOs at 8th grade and the closing of this gap at 6th grade may be associated with implementation effects and program coherence rather than the benefits of one instructional program over another.

Yet, how does this theory align with August, Goldenberg, & Rueda's (this volume) findings about the positive effects of reading instruction in students' primary languages on reading achievement in English? If reading instruction in students' primary languages leads to higher achievement in English reading, why do English learners at all grade levels not show test score trend lines consistent with the hypothesis that Prop 227 caused harm? Although we cannot

answer this question definitively, it is important to remember that, as August et al. (this volume) point out, the positive effects of primary language instruction are small to moderate in size. Furthermore, perhaps the positive effects of primary language instruction are also limited in duration, given the prevalence of transitional rather than maintenance models, and the fact that any positive effects in the early grades may be obscured by the powerful effects of school quality as students advance in their schooling careers. Maybe by the time students are in middle school, having experienced a coherent instructional program – bilingual or not – matters more for students’ academic achievement than the effects of bilingual education in the early elementary grades. Whatever the correct interpretation may be of our findings, it seems fairly clear that the sledgehammer approach to dismantling bilingual education in California disrupted any coherence that may have existed in the system, harmed students in the immediate term, and on average created no overall benefit. Draconian educational policy changes as embodied in Prop 227, it would appear, are best avoided, tempting as they might be for policymakers.

Given the many questions that remain, future research could seek further detailed information about what specific practices at the district, school, and classroom level are associated with educational success for English learners. Specifically, building on our findings, researchers could conduct in-depth case studies comparing specific districts in which the gap between ELs and EOs is closing to those in which it is widening. These case studies could seek to determine differences in districts’ practices possibly associated with achievement trends for language minority students but not captured by information in the state data system.

In conclusion, while definitive statements about the impact of Prop 227 remain elusive, it is clear that current and former English learners are not achieving the same levels of academic success as their peers who enter school already knowing English. While this achievement gap

appears to have closed slightly at some grade levels and widened slightly at others over the last five years, we do not see a clear association between the implementation of Prop 227 and consistent achievement gains for English learners relative to English-only students. If we aim to meet the challenge of closing this achievement gap and enabling all of California's more than six million public school students to experience academic success, we urgently need further research on effective policies and instructional practices for language minority students.

References

- August, D., Goldenberg, C. & Rueda, R. (This volume). Restrictive state language policies: Are they scientifically based?
- Bulter, Y.G., Orr, J.E., Gutierrez, M.B., & Hakuta, K. (Winter & Spring 2000). "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.
- California Secretary of State. (1998). "1998 California Primary Election Voter Information Guide/Ballot Pamphlet – English Language in Public Schools." Retrieved on January 28th, 2007 from <http://primary98.sos.ca.gov/VoterGuide/Propositions/227.htm>
- California State Department of Education. (February 5th, 2008a). "Report: Number of English Learners by Language, 2006-07." *Dataquest*. Retrieved on February 5th, 2008 from <http://data1.cde.ca.gov/dataquest/LEPbyLang1.asp?cChoice=LepbyLang1&cYear=2006-07&cLevel=State&cTopic=LC&myTimeFrame=S&submit1=Submit>
- California State Department of Education. (February 5th, 2008b). "Report: English Learners, Instructional Settings and Services, 2006-07." *Dataquest*. Retrieved on February 5th, 2008 from http://data1.cde.ca.gov/dataquest/EIP2_State.asp?RptYear=2006-07&RptType=ELPart2_1a
- California State Department of Education. (February 5th, 2008c). "Part II English Learner (EL) students, 1997-98, Number of English Learner Students receiving instructional services." *Dataquest*. Retrieved on February 5th, 2008 from http://data1.cde.ca.gov/dataquest/oEIPart2_1.asp?RptYear=1997-98&RptType=ELPart2_1a
- California State Department of Education. (n.d.). "Glossary - English Learner Students Receiving Instructional Services." *Dataquest*. Retrieved on January 30th, 2008 from http://dq.cde.ca.gov/dataquest/gls_EIPart2b.asp
- Castañeda vs. Pickard*. (1981). Retrieved on January 28th, 2008 from <http://www.stanford.edu/~kenro/LAU/IAPolicy/IA1bCastaneda.htm>
- Gándara, P. (Winter and Spring, 2000). "In the Aftermath of the Storm: English Learners in the Post-227 Era." *Bilingual Research Journal*, 24(1-2).
- Gándara, P., Maxwell-Jolly, J., Stritikus, T., Curry, J., Garcia, E., Asato, J., & Gutiérrez, K. (2000). *The initial effects of Proposition 227 on English learners*. Santa Barbara: University of California, Linguistic Minority Research Institute.
- Gándara, P., Rumberger, R., Maxwell-Jolly, J. and Callahan, R., (October 7th, 2003). English Learners in California Schools: Unequal resources, unequal outcomes. *Education Policy Analysis Archives*, 11(36). Retrieved January 27, 2008 from <http://epaa.asu.edu/epaa/v11n36/>

- Gold, N. (2006). *Successful Bilingual Schools: Six Effective Programs in California*. San Diego: San Diego County Office of Education.
- Grissom, J.B. (July 30, 2004). "Reclassification of English Learners." *Education Policy Analysis Archives*, 12(36). Retrieved January 27, 2008 from <http://epaa.asu.edu/epaa/v12n36/>
- Lau vs. Nichols*. (1974). Retrieved on January 28th, 2007 from <http://www.stanford.edu/~kenro/LAU/IAPolicy/IA1aLauvNichols.htm>
- Maxwell, Jolly, J. (2000). "Factors influencing implementation of mandated policy change: Proposition 227 in seven northern California school districts." *Bilingual Research Journal* 24(1-2).
- Parrish, T., Linqunti, R., Merickel, A., Quick, H. E., Laird, J., & Esra, P. (2002) *Effects of the Implementation of Proposition 227 on the Education of English Learners, K-12: Year 2 Report*. Palo Alto, CA: American Institutes for Research.
- Parrish, T., Perez, M., Merickel, A., & Linqunti, R. (2006). *Effects of the Implementation of Proposition 227 on the Education of English Learners, K-12: Findings from a Five-Year Evaluation*. Palo Alto, CA: American Institutes for Research.
- Williams, T., Hakuta, K., Haertel, E., et al (2007). *Similar English Learner Students, Different Results: Why Do Some Schools Do Better? A follow-up analysis based on a large-scale survey of California elementary schools serving low-income and EL students*. Mountain View, CA: EdSource.
- U.S. Department of Education, National Center for Education Statistics. (2006). [Public Elementary and Secondary Students, Staff, Schools, and School Districts: School Year 2003-04](#) (NCES 2006-307).

Figure 1: Grade Level Slices Included in Analysis

Cohort panel shows grade level by year of CST testing, marking the year when Proposition 227 was implemented, and the state testing program that was in place during each year. All students in cells above the dotted line entered school after Prop 227 was implemented. Note that all students in the 8th grade slice entered school before or during the year that Prop 227 took effect.

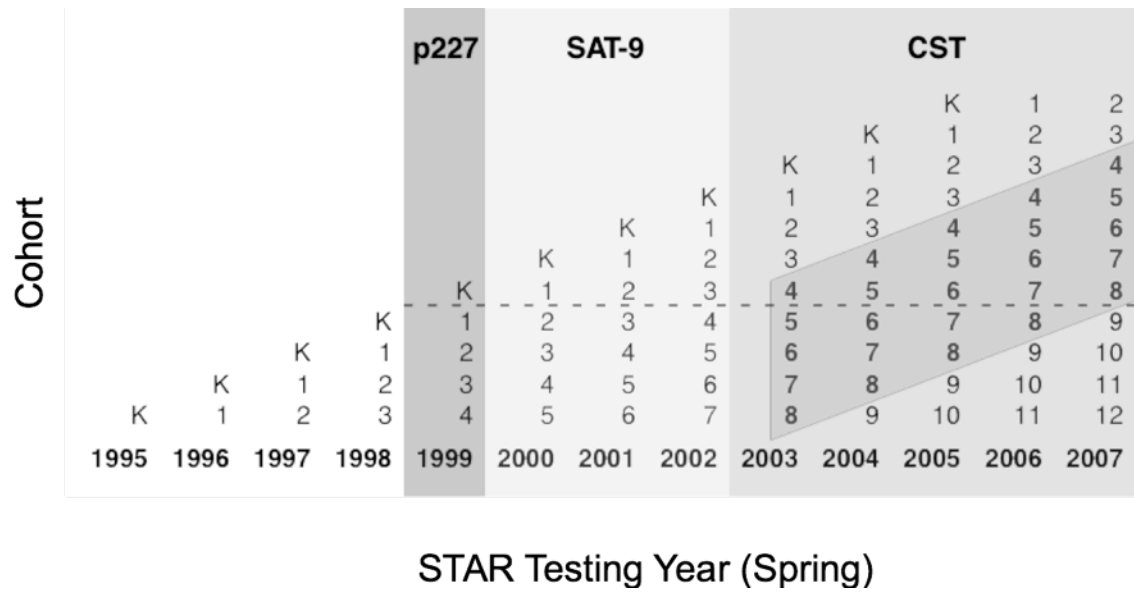


Figure 2: Predicted Results if Prop 227 Showed “Benefit” or “Harm” for English Learners
 Theoretical predictions for the available CST data by grade level are shown in the lower panels for Grade 6 and 8 for illustrative purposes, showing trend lines for EOs and ELs that would be parallel if no impact from Prop 227 occurred. Geometric wedges representing student cohorts are proportionate to the number of cohorts in each grade level slice that had pre- and post-Prop 227 experience, respectively. The impact of pre- and post-Prop 227 experience could either be positive (upward arrow) or negative (downward arrow), and would affect the slope for the ELs in either positive or negative directions, suggesting either that Prop 227 was associated with possible benefit or possible harm for ELs. It is assumed that Prop 227 did not have an effect on EO slopes.

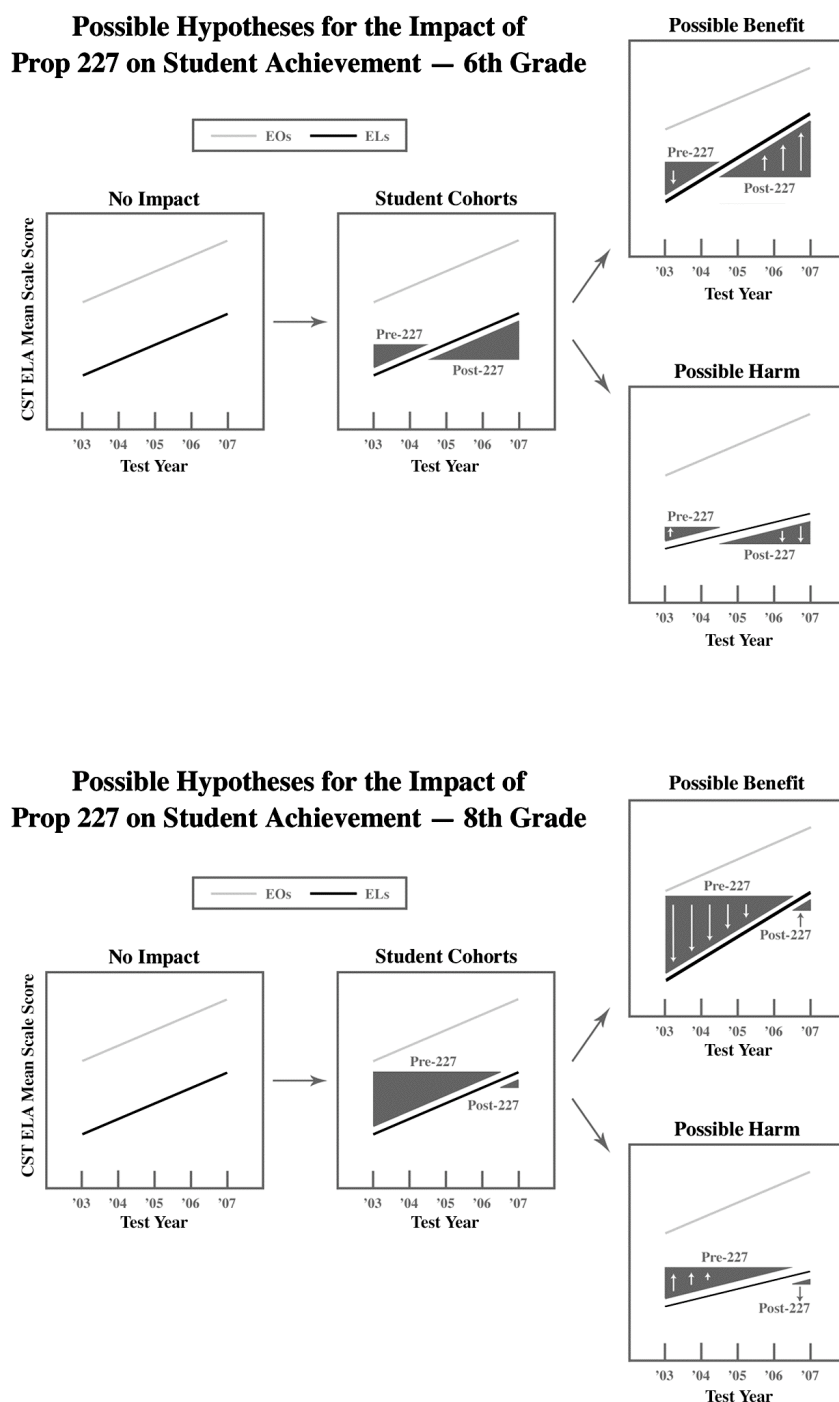


Figure 3

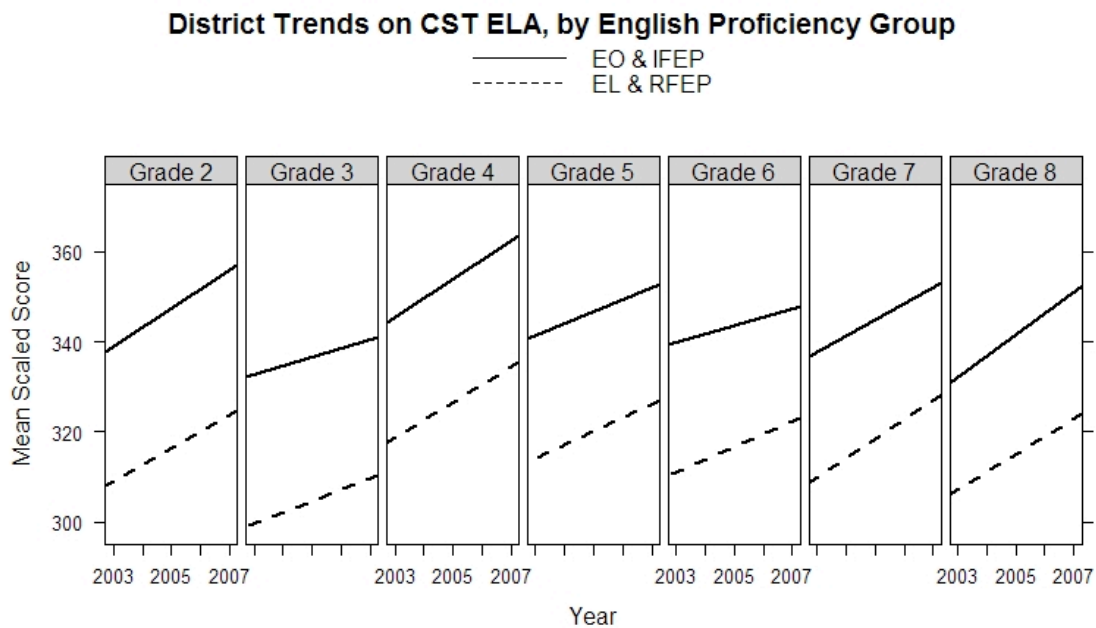


Table 1. District Level Estimates of CST ELA Slope for English Proficiency Groups

Grade	EL & RFEP		EO & IFEP		Corr.	difference in slope		t-test on difference ¹			
	Mean	Std. Dev.	Mean	Std. Dev.		Mean	Std. Dev.	Test Stat.	p.value ²	Lower ³	Upper ³
Grade 2	3.703	2.985	4.230	2.141	0.43	-0.527	2.831	-3.531	***	-0.821	-0.234
Grade 3	2.493	2.714	1.873	2.135	0.34	0.620	2.825	4.160	***	0.327	0.913
Grade 4	4.028	2.217	4.215	1.904	0.41	-0.187	2.258	-1.573		-0.422	0.047
Grade 5	2.980	2.285	2.716	1.701	0.32	0.264	2.370	2.113	*	0.018	0.510
Grade 6	2.760	2.435	1.860	1.775	0.35	0.900	2.456	6.943	***	0.645	1.155
Grade 7	4.203	2.547	3.645	2.186	0.48	0.558	2.424	4.364	***	0.307	0.810
Grade 8	3.805	2.522	4.688	2.076	0.47	-0.883	2.390	-7.007	***	-1.132	-0.636

Note: There are 359 observations for each combination of grade level and language proficiency group.

¹ Correlated samples t-test

² The probability of obtaining a result at least as extreme as observed, under the null hypothesis of equal means:
* = $p < .1$, ** = $p < .01$, *** = $p < .001$

³ 95% confidence interval for difference in slope within district

Figure 4

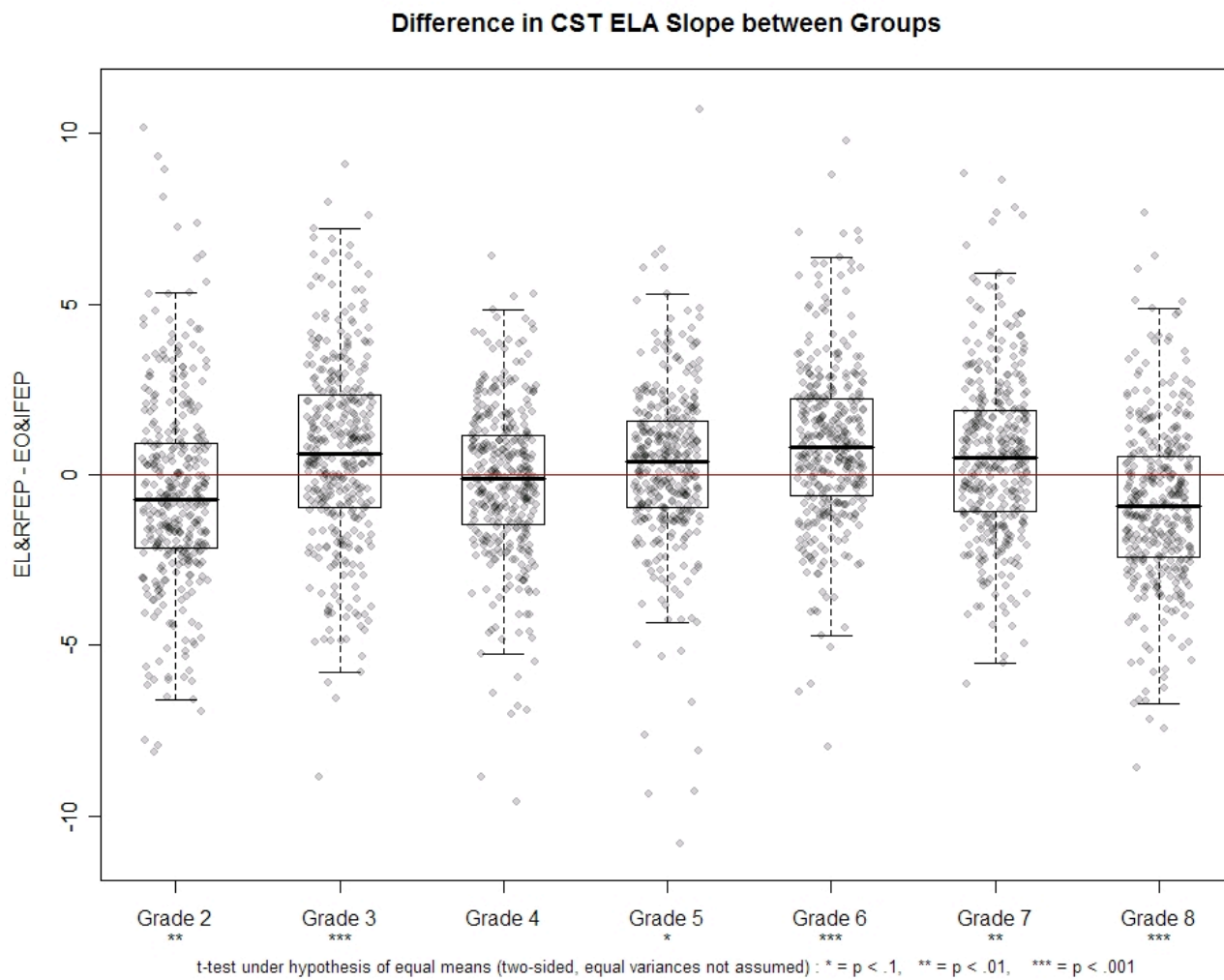


Figure 5

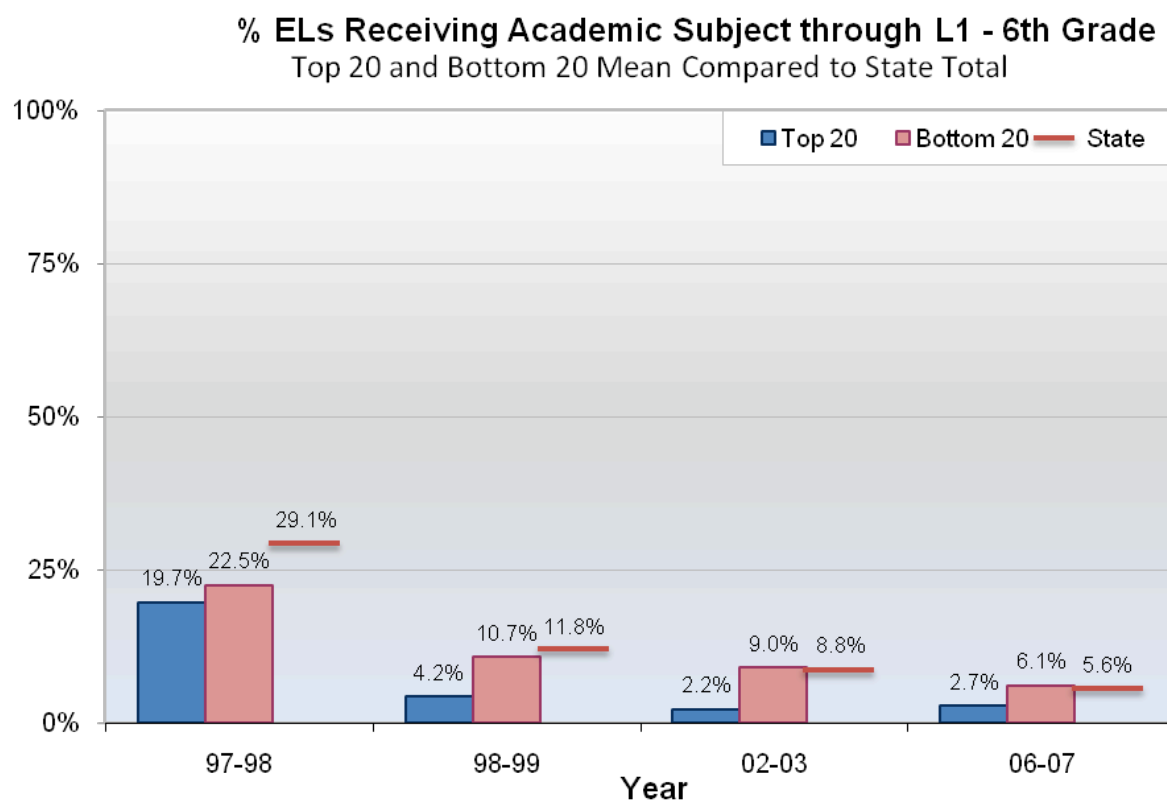


Table 2: Percentage of ELs Receiving Academic Subjects through L1 -Top 20/Bottom 20 Districts, 6th Grade

	Top 20			Bottom 20			sig. ^a
	mean %	mean %*	std. dev. %*	mean %	mean %*	std. dev. %*	
1997-1998	19.7%	0.696	0.736	22.5%	0.839	0.609	0.509
1998-1999	4.2%	0.208	0.393	10.7%	0.433	0.565	0.151
2002-2003	2.2%	0.164	0.263	9.0%	0.372	0.546	0.134
2006-2007	2.7%	0.172	0.301	6.1%	0.286	0.446	0.352

* District percents transformed using arcsin: $2(\arcsin(\sqrt{\%}))$

^a Significance (p-value) of two sample t-test testing for equality of transformed means

Figure 6

% ELs Receiving Academic Subjects through L1 - 8th Grade
 Top 20 and Bottom 20 Mean Compared to State Total

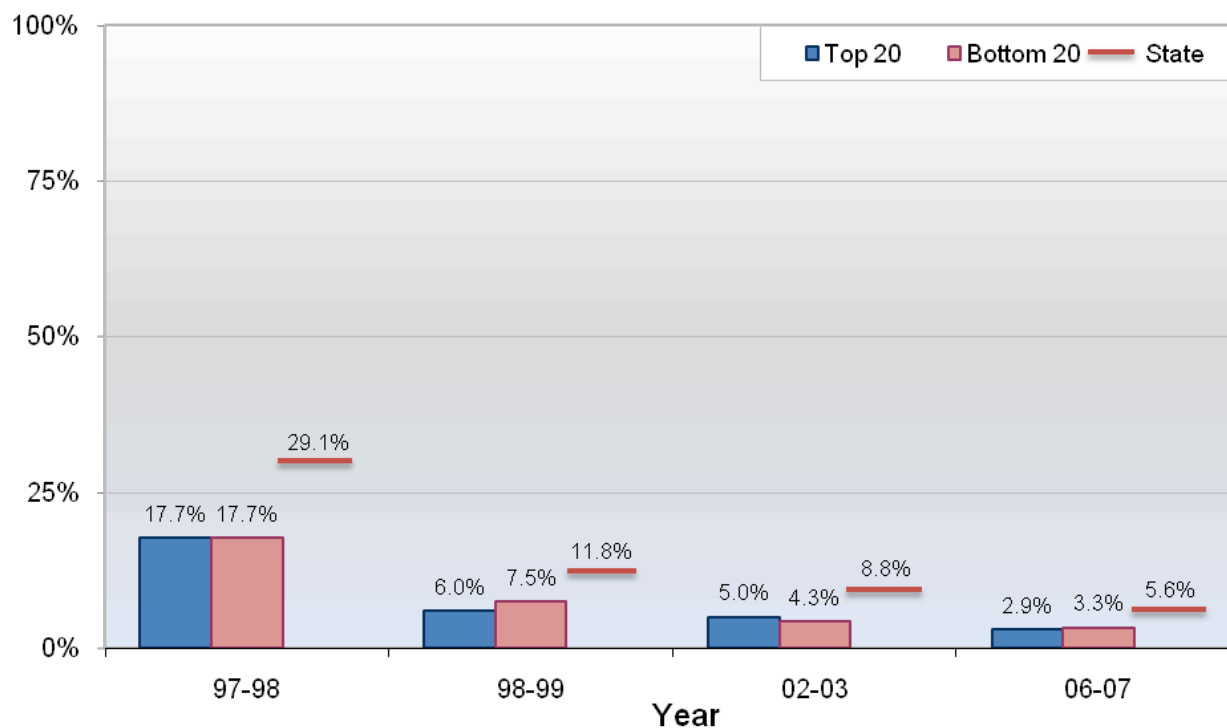


Table 3: Percentage of ELs Receiving Academic Subjects through L1 -Top 20/Bottom 20 Districts, 8th Grade

	Top 20			Bottom 20			sig. ^a
	mean %	mean %*	std. dev. %*	mean %	mean %*	std. dev. %*	
1997-1998	17.7%	0.711	0.573	17.7%	0.749	0.510	0.826
1998-1999	6.0%	0.295	0.432	7.5%	0.314	0.507	0.898
2002-2003	5.0%	0.257	0.415	4.3%	0.224	0.382	0.798
2006-2007	2.9%	0.147	0.337	3.3%	0.194	0.330	0.657

* District percents transformed using arcsin: $2(\arcsin(\sqrt{\%}))$

^a Significance (p-value) of two sample t-test testing for equality of transformed means

Figure 7

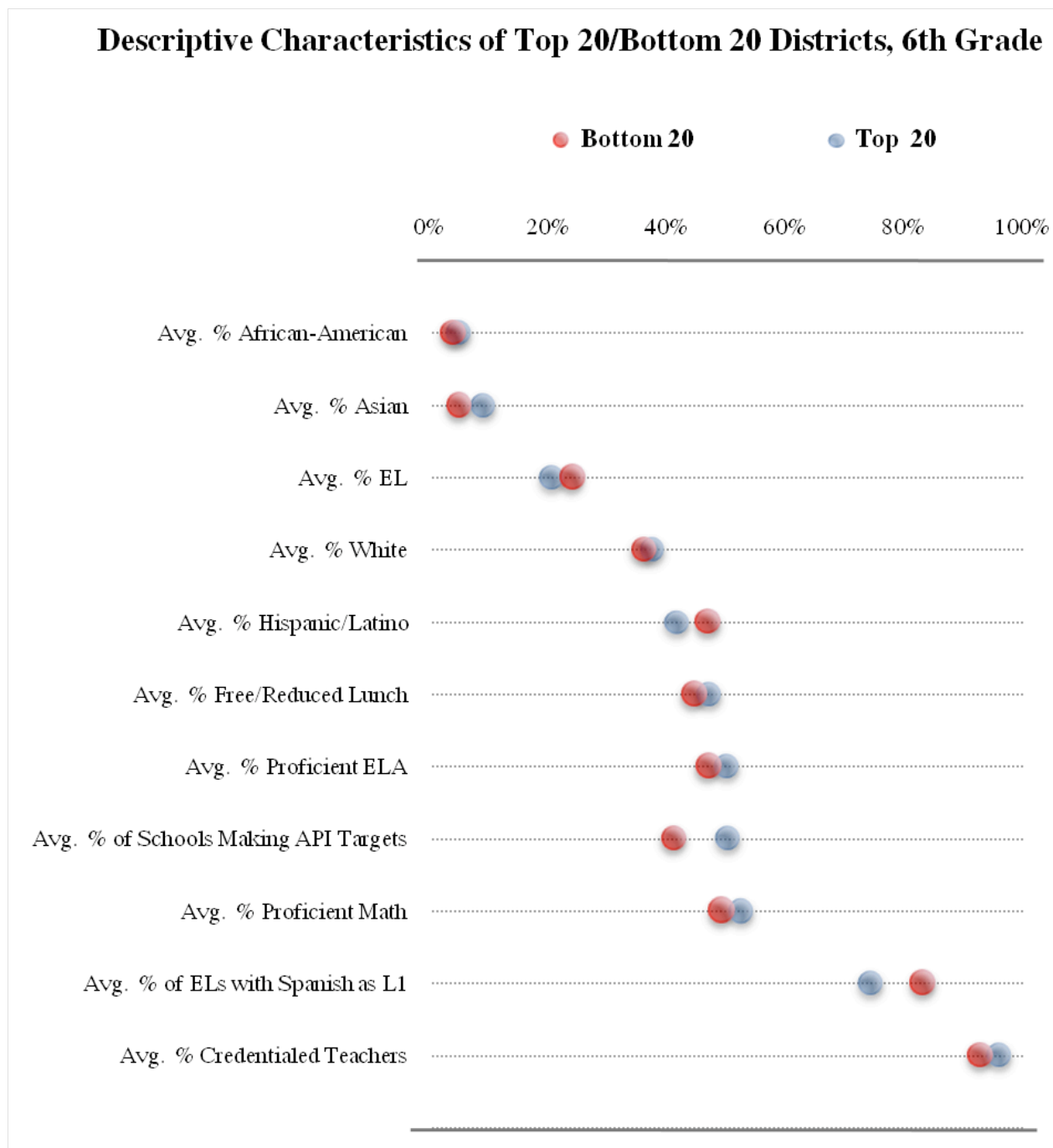


Figure 8

