

Race and Achievement Gaps in CAASPP Testing: A Decade of Data from LVUSD

Michael Bronshteyn

Calabasas High School, 22855 Mulholland Hwy, Calabasas, CA 91302, United States

ABSTRACT

This study analyzes racial disparities in educational outcomes. Using California Assessment of Student Performance and Progress (CAASPP) math and English testing data from the Las Virgenes Unified School District (LVUSD) spanning 10 years (2015–2024), it highlights the correlation between race and subsequent test scores. LVUSD is generally regarded as a high-performing district, yet inequities persist, illustrating how race influences standardized outcomes even in well-resourced contexts. The forces fueling inequities, resource distribution, implicit bias, and structural barriers are indicative of national trends despite the single-district focus of analysis. While the LVUSD case illustrates disparities that mirror national patterns, the findings are specific to one district and should be interpreted with caution when extrapolated to broader U.S. contexts. For example, African American students' math performance showed an R^2 of 0.689, meaning nearly 69% of the variance was associated with race, while White students' scores showed an R^2 of 0.404. In English, Hispanic students displayed the strongest correlation ($R^2 = 0.380$). ANOVA confirmed significant between-group differences ($F = 363.54$ for math, 163.29 for English; $p < 0.001$). Future research could compare results across additional districts and incorporate contextual factors such as school resources and home environments. The dataset, covering the 2015–2024 timeframe, excludes long-term outcomes and post-2024 policy reforms, which may also influence performance.

Keywords: Educational equity; Achievement gap; Standardized testing; Racial disparities; Structural barriers

INTRODUCTION

Educational equity has long been a chief principle of policy discussion in the United States, yet significant disparities persist, particularly among races. The

persistent disparities in educational outcomes, particularly among racial and ethnic groups, have been widely documented in the U.S. Research on educational inequality often highlights how systemic issues like socioeconomic level, discriminatory practices, and implicit biases disproportionately affect minority students.

Kearney and Levine discuss the interplay between income inequality, social mobility, and high school dropout rates, showcasing how systemic socioeconomic barriers disproportionately hinder minority students' educational trajectories (1). In parallel, Jones

Corresponding author: Michael Bronshteyn, E-mail: mbronshteyn070@gmail.com.

Copyright: © 2025 Michael Bronshteyn. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Accepted September 12, 2025

<https://doi.org/10.70251/HYJR2348.35236241>

et al. illustrate that historically marginalized groups and those from lower socioeconomic backgrounds are disproportionately affected in terms of academic achievement (2). This stems from a multitude of factors, including limited financial resources, unstable housing, underfunded schools, and reduced access to mental health services. These compounding barriers intensify educational disruptions and hinder long-term recovery for affected students.

Furthermore, Fryer Jr. and Levitt showcase that at kindergarten entry, the Black-White test score gap is largely explained by socioeconomic factors like income and parental education (3). However, the gap widens significantly during the first two years of school, even after controlling for these factors, suggesting that other influences, such as school quality or systemic biases, play a role. The study emphasizes that addressing socioeconomic variances alone is insufficient and calls for further research into structural and cultural factors impacting educational outcomes. Moreover, Spector highlights how Black students are inordinately subjected to suspensions and expulsions, which disrupt their learning and contribute to lower academic performance over time (4). Addressing inequitable disciplinary practices is essential to reducing the achievement gap and fostering more equitable educational outcomes.

Against this backdrop, standardized tests such as California's CAASPP play a pivotal role in assessing student performance and guiding policy reform. However, these assessments frequently illustrate levels of disparities in test scores by race, highlighting the need to discover the causes and consequences of these discrepancies. Current studies underscore how assessment can reproduce and intensify systemic inequalities. Au argues that high-stakes testing on standardized tests is "unequal by design," reinforcing structural disadvantage rather than serving as merit-neutral measures of ability (5). Similarly, Pierson, using Bayesian threshold tests to analyze COVID-19 screening, demonstrates how racial bias can materialize in seemingly objective assessment systems so that minority groups are disproportionately exposed to greater barriers to equal access (6). Cumulatively, the following findings suggest that standardized tests are not only diagnostic tools but also reflect and instantiate more pervasive racial disparities within American institutions.

Despite this extensive body of research, there is limited empirical work that utilizes a decade-long dataset, covering elementary, middle, and high

schools, at the district level. This study aims to fill that gap by analyzing racial disparities in CAASPP performance within LVUSD. In contrast to existing literature, this research focuses specifically on a single school district to examine localized patterns that otherwise would be overlooked in national datasets. By combining statistical analysis with contextual exploration of socioeconomic and institutional factors, this study seeks to provide practical insights for mitigating educational inequities.

METHODS AND MATERIALS

This study utilized data from LVUSD's CAASPP standardized testing, covering the years 2015 to 2024, excluding 2020 because of COVID-19, to investigate the relationship between race and academic performance in math and English. Statistical analysis to observe such performance included R-squared (RSQ), Confidence intervals, standard deviations, and single-factor ANOVA to assess variations in performance across racial groups.

Statistical tools were selected in this study based on their ability to quantify and test disparities across demographic groups. R^2 was utilized to determine how variance in test scores could be attributed to race, providing a direct measure of racial influence on academic performance. Confidence intervals were calculated to assess the reliability of the score averages, ensuring robustness in statistical conclusions. The single-factor ANOVA method was used due to its effectiveness in evaluating whether mean differences across racial groups were statistically significant, providing insight into group-level disparities.

The dataset consisted of district-level 2015-2024 (excluding 2020) CAASPP math and English scores. Scores were aggregated by year and stratified by racial group (African American, White, Hispanic, Asian). Data were not available at the individual level; therefore, the analysis is based on group-level findings for the district as a whole.

R^2 was calculated for each racial group individually to determine the proportion of variance in scores explained by race within a group's performance distribution. While R^2 is typically presented for an overall regression model, decomposing it by racial group here allowed a comparison of the extent to which race predicted outcomes for each group. This approach underscores that race's impact is not uniform but varies in strength across groups.

Data Collection

This study analyzed standardized CAASPP math and English scores from the Las Virgenes Unified School District (LVUSD), covering the years 2015–2024, excluding 2020 due to COVID-19 test cancellations. Scores were aggregated by racial group (African American, White, Hispanic, Asian) at the district level. The primary independent variable is race, while the dependent variables are math and English test scores.

To evaluate disparities, several statistical tools were applied in sequence. The first step was to calculate the coefficient of determination, or R^2 , which quantifies how much of the variance in scores can be explained by race. As shown in Equation (1):

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} \quad (1)$$

Next, the reliability of mean scores was assessed through confidence intervals. These were calculated using Equation (2):

$$CONF = t \cdot \left(\frac{s}{\sqrt{n}} \right) \quad (2)$$

To evaluate the spread of scores within each racial group, standard deviations were computed using Equation (3):

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n - 1}} \quad (3)$$

Finally, disparities between groups were tested for statistical significance using single-factor ANOVA. The procedure began with the calculation of between-group variance (Equation 4):

$$SS_B = \sum n_i (\bar{x}_i - \bar{x})^2 \quad (4)$$

and within-group variance (Equation 5):

$$SS_W = \sum (x_{ij} - \bar{x}_i)^2 \quad (5)$$

These values were then used to compute the F-statistic (Equation 6):

$$F = \frac{MS_B}{MS_W} \quad (6)$$

RSQ values for math scores were calculated for each racial group to measure the proportion of variance explained by race; with African American students having an RSQ of .0380, White students having an RSQ of .404, Hispanic students having an RSQ of

.008, and Asian students having an RSQ of .078. Confidence norms for math were recorded, including a value of 11.4, reflecting the reliability of the mean estimates. Standard deviations were measured, with math scores illustrating a value of 17.4. Lastly, ANOVA was measured, displaying a between-group variance of 3451.92, a within-group variance of 9.49, an F-statistic of 363.54, and a P - value of 0.00, demonstrating significant differences across racial groups.

For English scores, RSQ values were similarly calculated, with African American students showing an RSQ of .070, White students having an RSQ of 0.350, Hispanic students having an RSQ of 0.380, and Asian students having an RSQ of 0.082. Confidence norms for English were recorded, including a value of 8.26, reflecting the reliability of the mean estimates. Standard deviations were measured, with English scores illustrating a value of 12.6. Lastly, ANOVA was measured, displaying a between-group variance of 1753.6, a within-group variance of 10.73, an F-statistic of 163.29, and a P - value of 0.00, demonstrating significant differences across racial groups.

This study relied on statistical methodologies to bolster a rigorous analysis, leveraging confidence intervals and standard deviations to elucidate score reliability and variability with greater precision. Coefficients of determination (R^2 values) underscored the interplay between racial demographics and performance metrics, while the results of the ANOVA delineated the magnitude of observed disparities. These findings form a basis for scrutinizing deeper into the structural and systematic determinants influencing standardized testing performance.

RESULTS AND DISCUSSION

Math Performance Analysis

The analysis of math scores demonstrates clear disparities influenced by race (Figure 1). African American students experienced R^2 values of 0.689, meaning 68.9% of the variance in African American students' math scores can be explained by race (Equation 1). This strong relationship suggests that systemic issues like resource availability and implicit biases may disproportionately affect their performance. White students exhibited an R^2 value of 0.404, indicating that 40.4% of score variance is attributable to race. While still substantial, this lower value compared to African American students may reflect more equitable access to resources or differing systemic

barriers. Hispanic and Asian students exhibited R^2 values of 0.008 and 0.078, respectively, showing less significant correlations between academic achievement in terms of mathematical performance.

The confidence norm of 11.4 across all racial groups reflects the reliability of the mean scores (Equation 2). A lower confidence norm, such as 11.4, indicates higher precision in estimating the average score for a group. The value of 11.4 suggests moderate fidelity, suggesting that the variability in scores, likely due to external influences, must be considered in the examination of results. The standard deviation in this study measures the dispersion of scores within racial groups (Table 1). A standard deviation of

17.4 means that individual scores tend to deviate significantly from the mean (Equation 3). Such variability could stem from differences in access to resources, home environments, and faculty ability.

The Analysis of Variance (ANOVA) results in this experiment measure whether there are statistically significant differences in test scores - math and English - among the racial groups being examined. The between-

Table 1. ANOVA results for math CAASPP scores by race in LVUSD (2015–2024)

Statistic	Value	df	p-value
Between-group variance	3451.92	3	<0.001
Within-group variance	9.49	N-4	—
F-statistic	363.54	(3, N-4)	<0.001

Between-group and within-group variances, degrees of freedom, and significance levels are reported to illustrate the magnitude and reliability of disparities.

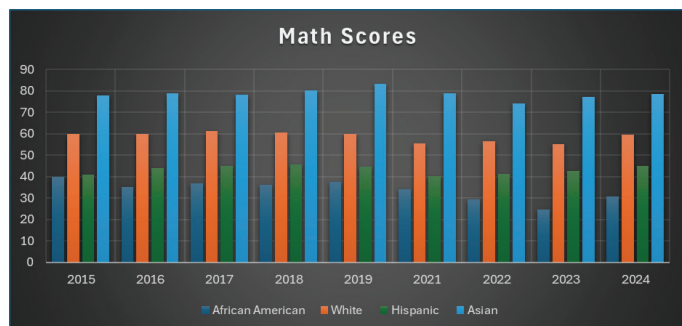


Figure 1. Math CAASPP scores (2015–2024) by race, showing achievement disparities across African American, White, Hispanic, and Asian students.

group variance quantifies the variability in math scores that can be attributed to differences between racial groups. A large value, 3451.92, indicates that racial differences amount to a significant portion of the overall score variance (Equation 4). The within-group variance represents variability within racial groups, such as differences in individual circumstances or school experiences. The relatively small value suggests that within-group differences are less prominent compared to between-group differences. The F-statistic measures the ratio of between-group variances to within-group variances. A high F-statistic like 363.54 means that differences between racial groups are much larger than differences within racial groups, strongly bolstering the notion that race influences math scores (Equation 6).

The P-value indicates the probability that the observed differences are due to chance. A P-value of 0.00 confirms that the disparities are statistically significant and unlikely to occur by chance.

English Performance Analysis

The analysis of English scores demonstrates clear disparities influenced by race (Figure 2). African American students experienced an R^2 value of 0.380, indicating that 38% of the variance in English scores can be explained by race (Equation 1). White students exhibited an R^2 value of 0.350, indicating that 35.0% of score variance is attributed to moderate racial influence, suggesting that systemic advantages might account for disparities for this group. Hispanic and Asian students exhibited R^2 values of 0.082 and 0.070, respectively, showcasing that race has a weaker correlation with English scores for these groups.

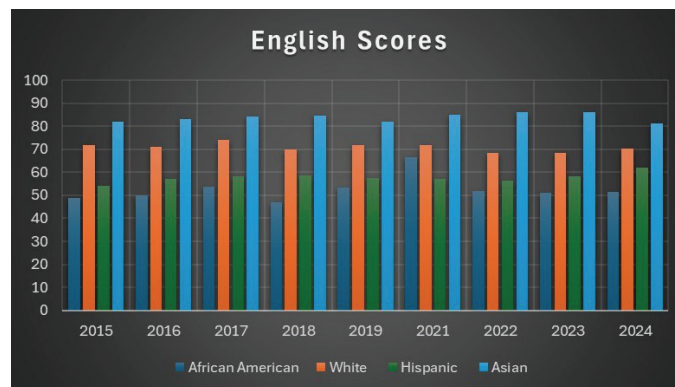


Figure 2. English CAASPP scores (2015–2024) by race, illustrating performance differences across African American, White, Hispanic, and Asian students.

The confidence norm of 8.26 across all racial groups reflects the reliability of the average English scores compared to math (11.4) (Equation 2). The lower value than math suggests less score variability, providing more precise estimates. With a smaller standard deviation than math (17.4), English scores exhibit less spread around the mean, suggesting more consistency in performance among racial groups (Equation 3).

The between-group variance, 1753.6, represents the variability in English scores attributed to differences between racial groups (Table 2). While smaller than the math variance, it still reveals a substantial racial influence. With the within-group variance at 10.73, it is slightly higher than the math, supporting that individual factors play a marginally larger role in English testing performance. The

F-statistic of 163.29 is lower than in math, indicating that racial differences, while significant, are less prominent in English testing (Equation 6). As with math, the P-value, 0.00, confirms that the disparities are statistically significant and not due to random variation.

CONCLUSION

This study highlights, through statistical analysis, significant differences in educational performance as measured by CAASPP test scores across racial groups within the Las Virgenes Unified School District (2015–2024).

For math scores, African American students exhibited the strongest correlation between race and performance ($R^2 = 0.689$), suggesting systemic inequities, such as resource disparities and implicit biases. However, R^2 is a correlation measure, and not a measure of causation; other potential explanations for this variation, such as differences in programming at the school, home, or unmeasured socioeconomic

levels, might also explain this variance. White students ($R^2 = 0.404$) showed a moderate correlation, while Hispanic ($R^2 = 0.008$) and Asian ($R^2 = 0.078$) students experienced weaker racial influences. English scores displayed different patterns, with Hispanic students showing the highest correlation to race ($R^2 = 0.380$), reflecting their potentially intricate challenges.

ANOVA results confirmed significant differences in performance between racial groups, with high between-group variances (3451.92 for math, 1753.6 for English) and F-statistics (363.54 for math, 163.29 for English), supported by P-values of 0.00. These findings showcase that disparities are not due to chance but are rooted in systematic factors.

While this study underscores the imperative to close gaps, some limitations need to be acknowledged. Firstly, the dataset was aggregated at the district level, precluding analysis by individual students or a cohort of grades. Secondly, only CAASPP math and English scores from 2015–2024 were examined, with other measures of measurement and longer-term metrics such as college enrollment or career attainment tests not considered. Thirdly, contextual factors such as home environment, access to mental health services, and more recent education reforms were not included, though they can influence test outcomes apart from systemic inequalities. These limitations indicate avenues for future research: (1) employing student-level longitudinal data to monitor within-group variation; (2) expanding to additional districts and states to address generalizability; (3) examining additional measures of performance other than CAASPP, including social-emotional measures; and (4) intersecting socioeconomic and institutional controls, such as funding structures and discipline policies, to produce a richer portrait of educational inequity.

Given the findings of this paper, potential policy measures to mitigate such profound disparities could entail reallocating wealth from affluent districts to historically underfunded schools serving predominantly minority populations. Progressive tax funding, reallocating wealth from affluent districts to historically underfunded schools, could help reduce disparities. While this study underscores the necessity to intervene with such inequities, it also highlights limitations, such as the lack of examination of home environments, mental health access, and recent educational reforms. Future research should expand to other districts, evaluate more factors contributing to academic performance, and consider more forms of testing besides CAASPP.

Table 2. ANOVA results for English CAASPP scores by race in LVUSD (2015–2024)

Statistic	Value	df	p-value
Between-group variance	1753.6	3	<0.001
Within-group variance	10.73	N-4	—
F-statistic	163.29	(3, N-4)	<0.001s

Statistical values highlight significant differences in mean performance across racial groups.

FUNDING

The author received no funding sources that supported the production of this paper.

CONFLICT OF INTEREST

The author declares there are no conflicts of interest regarding the publication of this study.

REFERENCES

1. Kearney MS, Levine PB. *Income inequality, social mobility, and the decision to drop out of high school*. National Bureau of Economic Research. Available from: https://www.nber.org/system/files/working_papers/w25254/w25254.pdf (accessed on 2025-04-08)
2. Jones TM, Fleming C, Williford A. Research and Evaluation Team of Seattle Public Schools. Racial Equity in Academic Success: The Role of School Climate and Social Emotional Learning. *Child Youth Serv Rev*. 2020 Dec; 119: 105623. <https://doi.org/10.1016/j.chilyouth.2020.105623>
3. Fryer RG, Levitt SD. Understanding the Black-White test score gap in the first two years of school. Price Theory at Chicago. Available from: <https://pricetheory.uchicago.edu/levitt/Papers/FryerLevittUnderstandingTheBlack2004.pdf> (accessed on 2025-5-8)
4. Racial disparities in school discipline are linked to the achievement gap between Black and white students nationwide, according to Stanford-led study. Available from: <https://ed.stanford.edu/news/racial-disparities-school-discipline-are-linked-achievement-gap-between-black-and-white> (accessed on 2025-04-08).
5. Au W. *Unequal by design: High-stakes testing and the standardization of inequality*. Routledge, 2022. ISBN: 978-1-032-04523-0, 1-216. <https://doi.org/10.4324/9781003005179-1>
6. Pierson E. Assessing racial inequality in COVID-19 testing with Bayesian threshold tests. Available from: <https://arxiv.org/abs/2011.01179> (accessed on 2025-05-08)