

## Making Students Visible: Comparing Different Student Subgroup Sizes for Accountability

CORE-PACE RESEARCH PARTNERSHIP

May 10, 2016

Heather Hough ([hjough@stanford.edu](mailto:hjough@stanford.edu))  
Joe Witte ([joe.witte@stanford.edu](mailto:joe.witte@stanford.edu))

### Summary and policy implications

With the passage of the Every Student Succeeds Act (ESSA) of 2015, California state policymakers are tasked with determining the subgroup threshold for school-level reporting. To inform this decision, this policy brief explores the implications of utilizing various subgroup sizes using data from the CORE Districts. We find that:

- A substantially higher percentage of student data is reported at smaller subgroup sizes. For example, when the subgroup size is reduced to 20+ from 100+, *six times* as many schools report results for African-American students.
- At a subgroup size of 20+, approximately ten times as many schools report results for *all student subgroups* than at a subgroup size of 100+.
- The lowest performing racial/ethnic subgroup in the school is often excluded from schools' reporting at the higher subgroup size. In 47 percent of the schools, results for the school's lowest performing racial/ethnic group are only reported when the subgroup size is reduced from 100+ to 20+.
- The reduction in subgroup size substantially changes school rankings; at the 20+ threshold, most schools' weighted test scores are lower than at the 100+ threshold. However, the schools identified in the bottom 5 percent of schools tend to be the same regardless of subgroup threshold.
- The 20+ subgroup size presents clear advantages in terms of the number of students represented, particularly in making historically underserved student populations visible. This increased information about subgroup performance may better support a continuous improvement framework, helping inform and empower educators to improve student outcomes.

## Purpose

With the passage of the Every Student Succeeds Act (ESSA) of 2015, California state policymakers must make many technical decisions about the state-wide accountability system. One of these decisions is the subgroup size required for school-level reporting. ESSA stipulates that states must break down the results of student performance on state assessments by subgroups, including major racial and ethnic groups, economically disadvantaged students, children with disabilities, English proficiency status, gender, and migrant status at the state, district and school level, but the determination of subgroup size is left to the states to decide.

Under NCLB in California, the subgroup size was set to 100+, which means that if there were fewer than 100 students in any number of subgroups – such as Hispanic or African-American students, English learners, or students with disabilities – then the assessment results from their unique subgroup were not reported at the school level. A smaller subgroup size could ensure that the performance of more students is reported.

In making decisions about how to define the minimum subgroup size, we hope the state can learn from the CORE Districts, who have developed a robust measurement system that represents nearly a million of California's students. In the School Quality Improvement Index (SQII), the CORE Districts utilize a subgroup size of 20+ instead of 100+ with the intention of including many more youth in the SQII than were previously part of the school accountability system.

To support the conversation as the State Board of Education considers which subgroup size to employ in the state-wide accountability system, PACE has conducted an analysis of CORE's SQII data to investigate how many more students are included when the subgroup size is reduced from 100+ to 20+ and how this reduction impacts school performance rankings.

## Approach

The analyses here are based on data from Fresno, Long Beach, Los Angeles, Oakland, San Francisco, and Santa Ana Unified School Districts. The analysis includes 1,030 schools, representing 12 percent of California's student population.<sup>1</sup> While the SQII represents a multiple-metric approach to measuring school performance encompassing both academic and social emotional factors, in this analysis we focus only on mathematics performance as measured by the SBAC in 2015.<sup>2</sup>

---

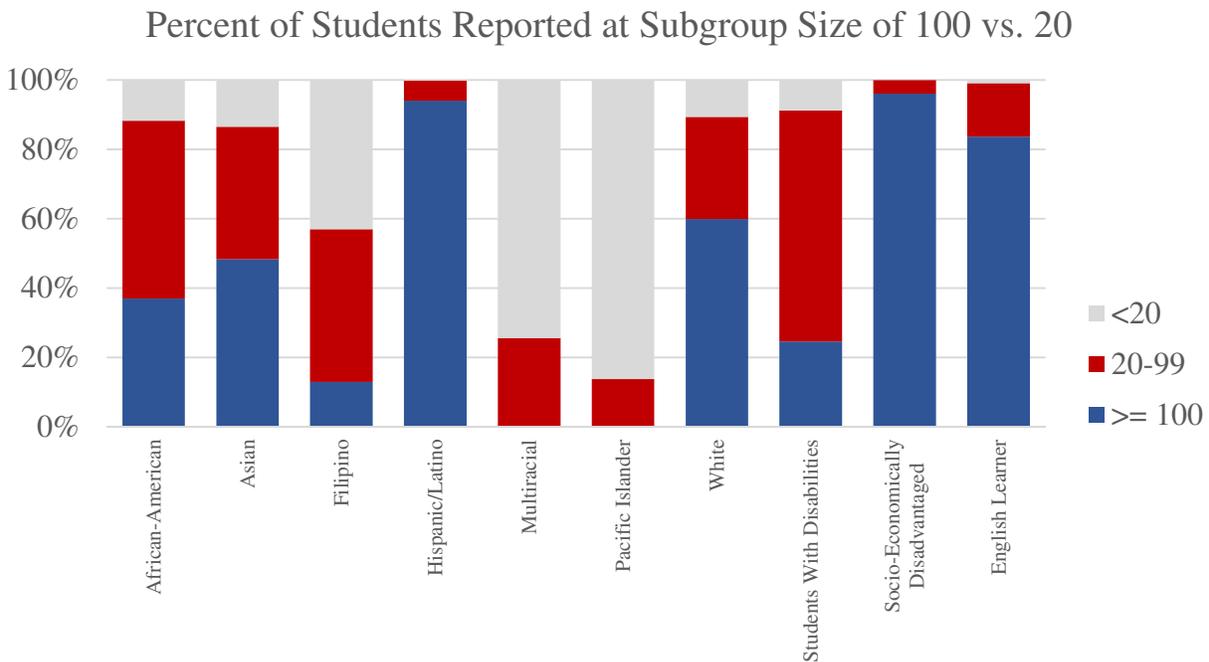
<sup>1</sup> There are 729 elementary schools, 149 middle schools, and 152 high schools represented in the CORE districts' data used in this analysis. This analysis does not include non-traditional span schools.

<sup>2</sup> In the analyses conducted for this brief, ELA and mathematics results are close to identical, so we only report mathematics results for simplicity.

## Student subgroups reported at subgroup size of 20 vs. 100

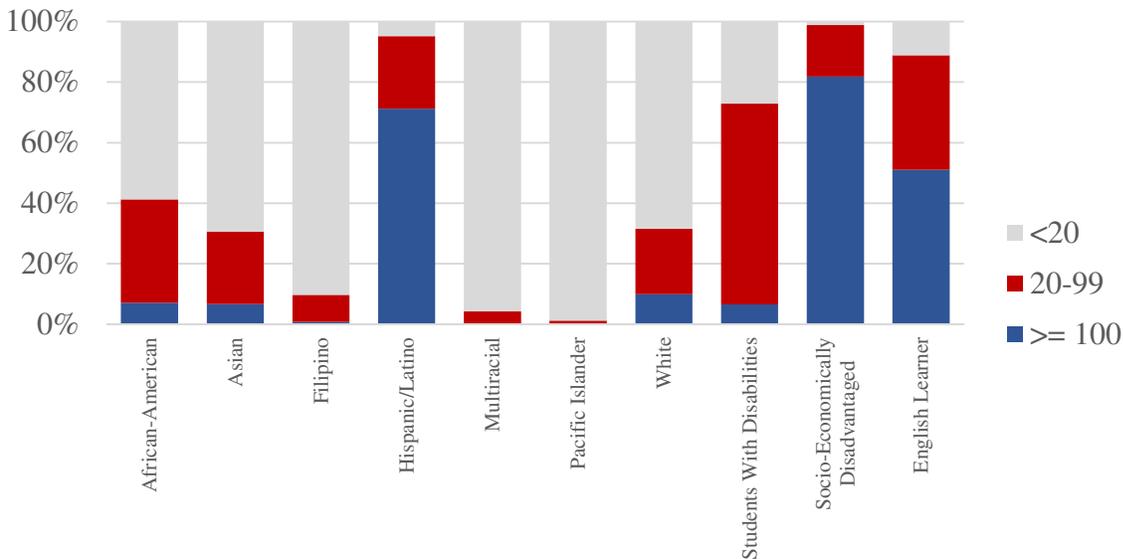
The most basic questions in investigating the impact of various subgroup sizes are 1) how many students are included, and 2) how many schools have visible subgroups at different size specifications?

The choice between setting minimum subgroup size at 20+ or 100+ has significant implications for how many students are counted at the school-level. As shown in the chart below, only 37 percent of African American students' math scores are reported at the school-level when the subgroup size is 100+, but this number reaches 88 percent when the subgroup size is decreased to 20+. This effect is even more dramatic for students with disabilities: only 25 percent of students with disabilities are reported at the school-level when the subgroup size is 100+, but this number reaches 92 percent if the subgroup size is decreased to 20+.



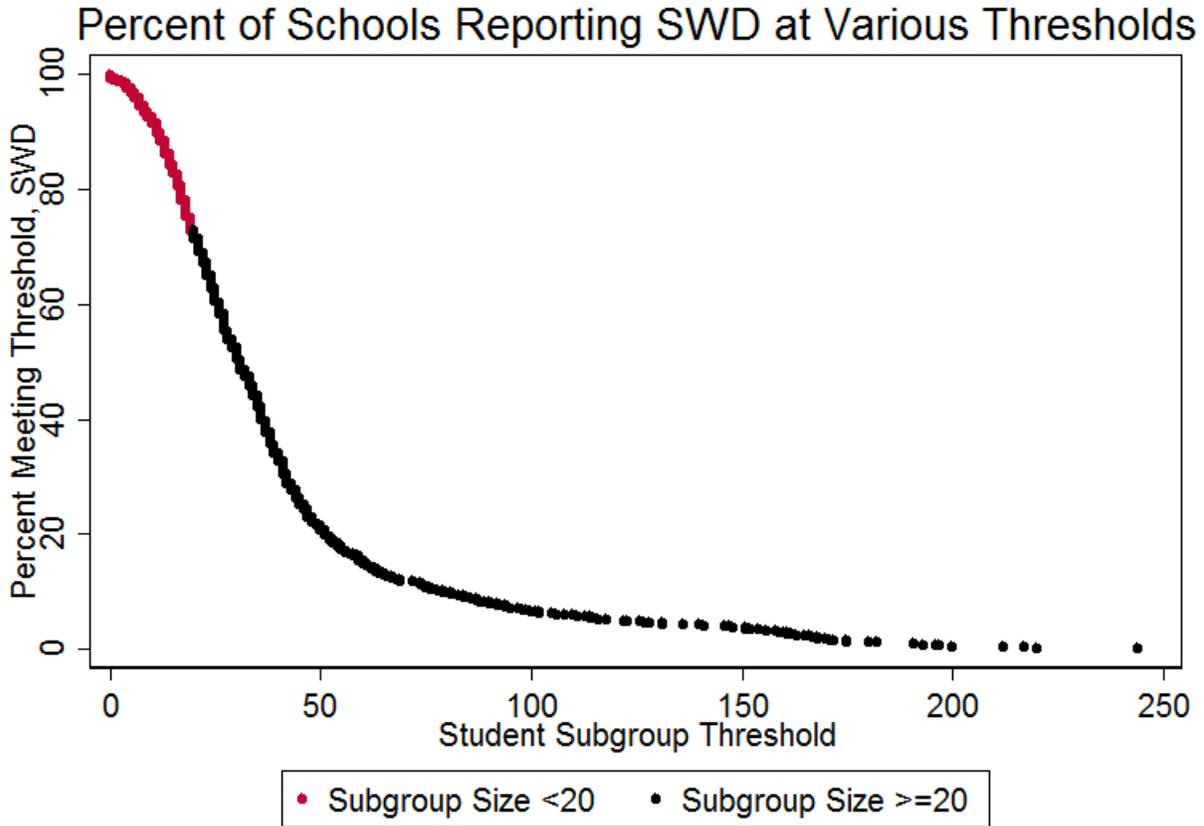
The difference in the percentage of *schools* reporting subgroups at the 100+ vs. 20+ thresholds is even more dramatic. Returning to the example of African-American students discussed above, only 7 percent of schools report an African American subgroup when the subgroup size is 100+, but this number reaches 41 percent when the subgroup size is decreased to 20+. This means that when the subgroup size decreases to 20+, 351 schools now report the math results of African American students that did not before. Effects are similarly large across multiple student subgroups – for instance, no school examined has a high enough percentage of Multiracial and Pacific Islander students at a 100+ threshold to display subgroup data, although both of those subgroups are represented at the 20+ threshold.

Percent of Schools Reporting Subgroups at Threshold of  
 100 vs. 20



To further illustrate the impact of various subgroup thresholds, the below chart shows the percent of schools that meet a given student subgroup threshold for students with disabilities (SWD). At a subgroup size of 0, 100 percent of schools meet the threshold. As subgroup size increases, the percentage of schools that meet that threshold rapidly decreases: 73 percent of the schools meet the threshold at 20+ subgroup size, 52 percent meet the threshold at 30+ subgroup size, 21 percent meet the threshold at 50+ subgroup size, and only 7 percent meet the threshold at 100+ subgroup size.<sup>3</sup> Other low-representation subgroups show similar patterns.

<sup>3</sup> It should be noted that starting from a 20+ subgroup size, a substantial amount of student data is excluded with almost any increase in subgroup size. For instance, between a 20+ and 50+ subgroup size, we can expect a 1.86 percent decrease in the percentage of schools meeting the threshold for inclusion with every 1-student increase in the subgroup threshold.



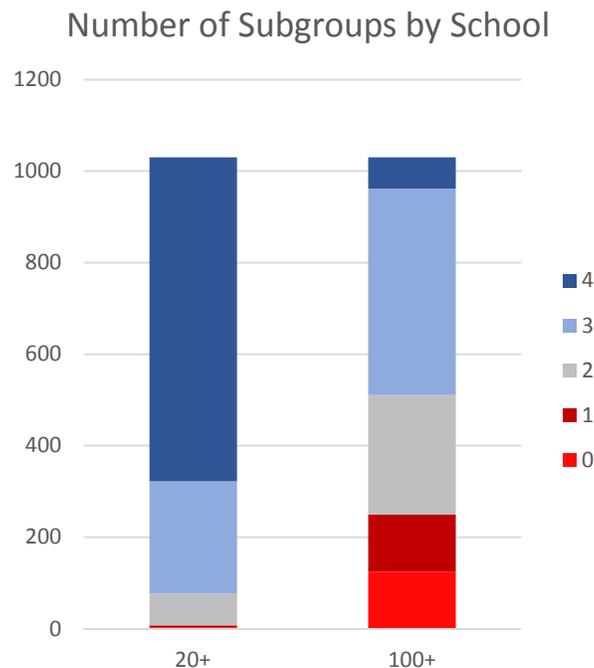
Taken together, these analyses show that a higher subgroup size masks the performance of student subgroups within many schools. If ESSA specifically requires the reporting of student achievement by subgroup, a high subgroup threshold effectively counteracts this requirement by failing to highlight the performance of small, but critical, groups of students, such as African American students or students with disabilities.

### **Inclusion of subgroups for accountability at threshold of 20 vs. 100**

In the CORE SQII, each metric is weighted to give equal importance to the “all students” group and the subgroups. Specifically, this means that for any given indicator, “all students” count for 50 percent of the measure, and the subgroups make up the remaining 50 percent, with 12.5 percent each allocated to the “lowest performing racial/ethnic group” (LPRG), disadvantaged students, students with disabilities, and English learners.

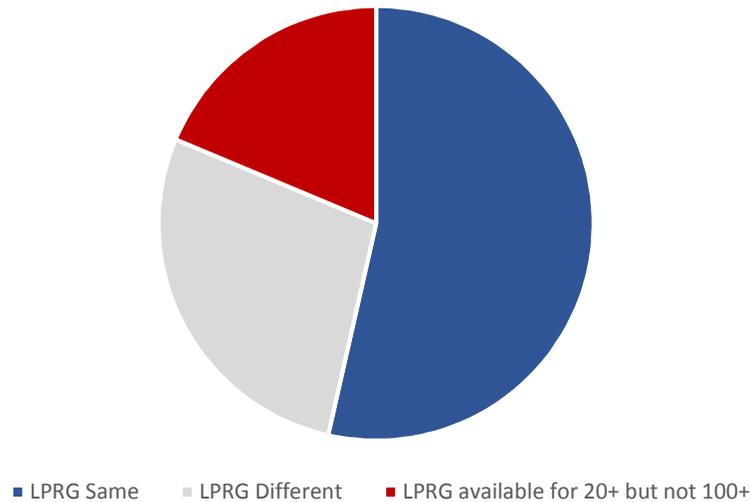
We see from these data that there is a significant difference in the number of subgroups reported by schools at a threshold of 20+ vs. 100+. As shown below, changing the subgroup size from

100+ to 20+ increases the count of schools with all four math index subgroups from 68 (7 percent) to 708 (69 percent). Moreover, at a subgroup size of 100+, 125 schools (12 percent) do not report *any* student subgroups. The combined effect of this loss of data at subgroup size 100+ is that most schools lose at least one index subgroup that would have been factored into the calculation at subgroup size 20+.



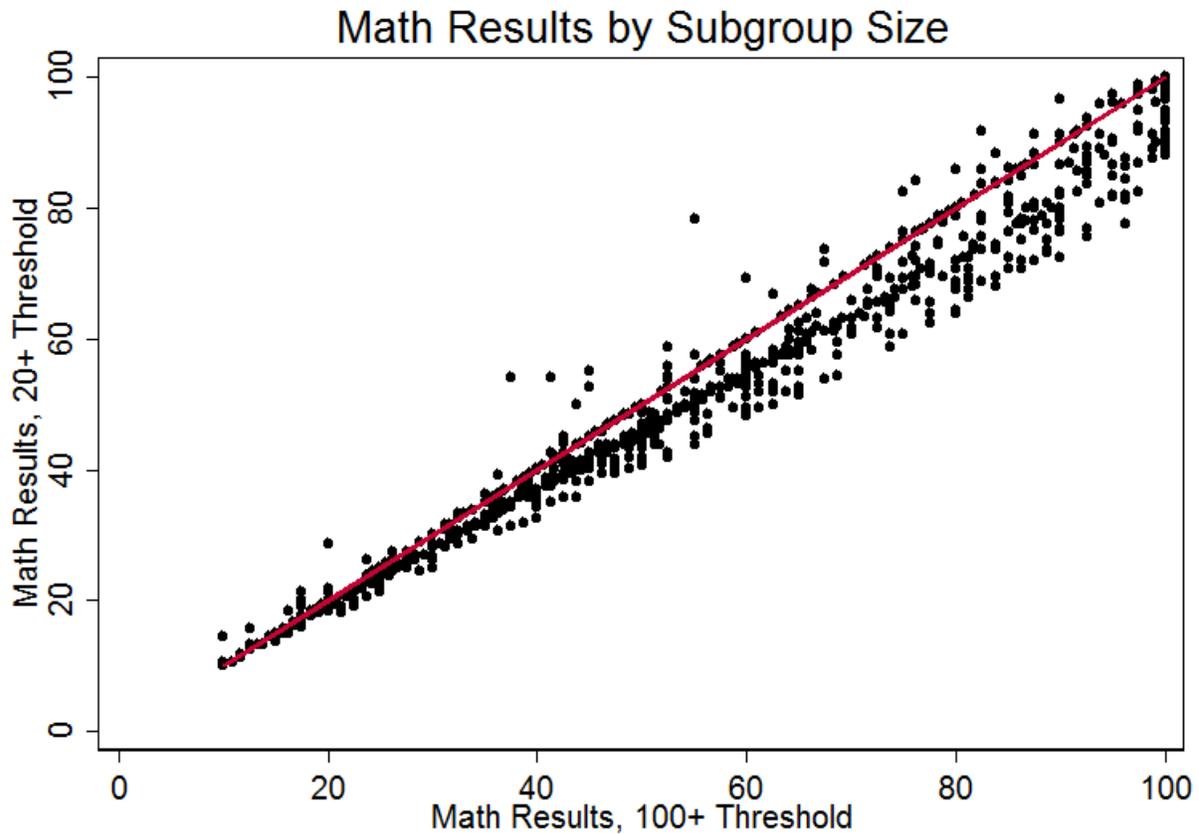
In addition, the identification of the lowest performing racial/ethnic group (LPRG) changes significantly depending on what subgroup size threshold is employed. If the subgroup size is increased from 20+ to 100+, 550 (54 percent) of schools retain the same LPRG, 285 (28 percent) change LPRG, and 192 (19 percent) now have insufficient data to determine LPRG. The student subgroup most affected by this change is African-American students – at a subgroup size of 20+, 333 schools report African-American students as their LPRG, but at a subgroup size of 100+, this number drops to 70. Additionally, while Pacific Islander, Multiracial, and Filipino students appear as the LPRG in some CORE schools at the 20+ threshold, there are not enough students in these subgroups for them to be included at a 100+ threshold. The changing of the LPRG when subgroup size is decreased indicates that the schools’ most vulnerable populations are exactly the students masked by the higher subgroup size.

### Changes in Identification of Lowest Performing Racial/Ethnic Group at Subgroup Size of 20+ vs. 100+



### School rankings at subgroup size of 20 vs. 100

Another important question is how the subgroup threshold changes the relative ranking of schools. The chart below shows the relationship between school-level average math scores when subgroups are reported at 100+ compared to 20+. Results above the line indicate that subgroups revealed at 20+ are performing higher than the subgroups reported at a threshold of 100+, and results below the line mean that subgroups revealed at 20+ are performing lower. For the vast majority of schools, mathematics performance is reported as lower when the subgroup size is reduced, indicating that the smaller subgroups that are made visible at 20+ that were not visible at 100+ contain lower performing students within the schools. This increased information about school performance indicates that a 20+ subgroup threshold may better support a continuous improvement framework, where improved levels of information about student subgroup performance can help inform and empower educators interested in supporting those populations of students more effectively.



It is important to note that the identification of the bottom 5 percent of schools for ESSA-mandated intervention does not appear to be substantially impacted by the choice of subgroup threshold – the lowest-performing schools tend to perform very poorly on mathematics assessments consistently across subgroups. For example, schools at this minimal performance level have extremely low rates of students meeting or exceeding standards: only 6 percent at both the 20+ and 100+ subgroup threshold.