

Identification of Comprehensive Support and Improvement Schools in California

The Influence of Selected Measures and System Design

Drew Atchison

Katie Hyland

Umut Ozek

Kerstin Le Floch

Steven Hurlburt

Institute of Education Sciences Award No: R305A200254

JULY 2023



Advancing Evidence.
Improving Lives.

Contents

- Executive Summary..... 1
 - Key Findings 1
- Background and Policy Context 3
 - Overview of California’s Accountability System 4
 - Situating California’s Approach Within the Broader Landscape of ESSA Accountability Systems 6
- Objectives and Research Questions 7
- Method 9
 - Data and Sample 9
 - Analytic Approach..... 9
- Results..... 11
 - RQ 1: Indicators Evaluating School Performance in Practice 11
 - RQ 2: Correlations Among Performance Measures and Between Performance Measures and School Demographics..... 14
 - RQ 3: Year-to-Year Consistency of Performance Indicators..... 22
 - RQ 4: Influence of Performance Indicators 24
- Discussion..... 28
- References 30
- Appendix A. Supplemental Exhibits 31

Exhibits

Exhibit 1. Every Student Succeeds Act (ESSA) Indicator Performance Categories and California’s School Performance Indicators.....	4
Exhibit 2. Definitions of Color Ratings for the English Language Arts (ELA) Performance Indicator	6
Exhibit 3. Percentage of All Public Schools and Comprehensive Support and Improvement (CSI) Schools in California by Number of Indicators Rated in 2019–20, Separately for Each Grade Level	12
Exhibit 4. Percentage of All Public Schools and Comprehensive Support and Improvement (CSI) Schools in California With a Given Color Rating for Each Indicator Rated in 2019–20, Separately for Each Grade Level.....	13
Exhibit 5. Correlations Among Measures of Performance Indicators for Elementary and Middle Schools.....	15
Exhibit 6. Correlations Among Measures of Performance Indicators for High Schools.....	16
Exhibit 7. Variance in Each Performance Indicator Explained by Measures of the Other Performance Indicators	17
Exhibit 8. Correlations Between School Performance Indicators and Measures of School Demographic Composition	18
Exhibit 9. Distribution of Measures of Status and Change Across Schools in the Highest and Lowest Poverty Quartiles, Separately for Each Performance Indicator	20
Exhibit 10. Color Ratings for Schools in the Highest and Lowest Poverty Quartiles, Separately for Each Performance Indicator.....	22
Exhibit 11. English Language Arts (ELA) Performance Status, Change, and Color Rating on the 2019 Dashboard by the ELA Performance Status, Change, and Color Rating in the 2018 Dashboard.....	24
Exhibit 12. Percentage of Schools That Would be Reclassified if a Given Performance Indicator Were Dropped From California’s Accountability System, Separately for All Schools and Each School Level.....	26
Exhibit 13. Number of Comprehensive Support and Improvement (CSI) Schools That Would Have Avoided CSI Designation With a Higher Color Rating, Separately for Each Performance Indicator	27
Exhibit A1. Variance in Measures of School Performance Indicators Explained by Student and School Characteristics	31

Exhibit A2. Status, Change, and Color Rating on the 2019 Dashboard by the Status,
Change, and Color Rating in the 2018 Dashboard for English Language Arts (ELA), Math,
Suspensions, and Chronic Absenteeism 32

Executive Summary

For the past quarter century, federal law has required states to measure school performance, identify the lowest performing schools, and provide support to those schools. Such systems have invited critiques—that they overemphasize student achievement in reading and mathematics, exclude nonacademic features of school performance, and rely on measures closely associated with student characteristics. The Every Student Succeeds Act (ESSA) of 2015 ushered in a new era of accountability, enabling states to design complex accountability systems containing multiple measures intended to provide a more holistic picture of a school’s overall performance. Based on these evaluations of school performance, states are required to identify the lowest performing 5% of Title I schools for comprehensive support and improvement (CSI).

ESSA extends wide latitude to states regarding the selection, weighting, and aggregation of various measures to identify low-performing schools. These state policy decisions have the potential to influence which schools are identified for support. These decisions also have implications for how school and district leaders respond to accountability designations, and ultimately, how they may influence student outcomes. In this report, we examine how specific design choices regarding California’s ESSA accountability system influence the set of schools identified for CSI.

Key Findings

To evaluate school performance and identify CSI schools, California assigns ratings to schools on seven possible indicators: (a) math performance, (b) English language arts (ELA) performance, (c) suspension rate, (d) chronic absenteeism (elementary and middle schools only), (e) graduation rate (high schools only), (f) college and career readiness (high schools only), and (g) English learner progress (ELP).¹ **In practice, however, CSI schools were held accountable for fewer indicators than what was typical of all schools. Thus, CSI schools were often identified on the basis of a partial portrayal of their performance. In particular, both ELA and math performance were much more likely to be unrated in CSI schools.**

State accountability systems include multiple—ideally complementary—measures, each of which should provide information on a different facet of school performance. However, if these

¹ For more details regarding the underlying data and how the indicators are measured, see the *2019 California School Dashboard Technical Guide* (California Department of Education, 2019). Not all schools are rated on every indicator due to various factors. Some indicators, such as graduation rates, are grade-specific and may not apply to all schools. Schools receive a rating for an indicator only if they have sufficient numbers of students contributing to the calculation of performance for that specific indicator.

measures are strongly correlated with each other, they provide little new information. Indeed, in California, **we found strong, positive correlations between the status measures, whereas correlations among the change measures of performance were weak or nonexistent.² In addition, among the status measures, suspension rate and ELP were the most weakly correlated with other measures, suggesting that they provide more distinctive information about schools compared with other measures.**

School accountability systems are intended to measure school performance, not simply to reflect the features of the student population. Measures that are strongly correlated with school demographics may be seen as unfair—identifying schools on the basis of the characteristics of the students they serve rather than on how well they serve those students. **In California, the status measures of school performance were moderately to strongly correlated with school demographics. In contrast, the change measures of school performance were virtually uncorrelated with student demographics, thus putting schools on a more even playing field.**

In addition to the fairness of accountability indicators, the reliability of these measures is also an important consideration. Fluctuations in measurement from year to year could signal noise or randomness; consistency over time suggests that the measure is stable and not random. Of course, one would not want to see perfectly stable measures; very high correlations would suggest that schools had failed to improve. **The California accountability system showed strong year-to-year consistency in the measurement of performance statuses. In contrast, the change ratings fluctuated substantially from year to year.**

Within a multiple-measures accountability system, each performance measure should contribute information to determine a school's overall performance and have some amount of influence in determining which schools ultimately become CSI schools. However, some measures may be more influential in determining school performance if they are distinct from others or are rated more often. **In California, the suspension rate indicator played a larger role than the other performance indicators in determining which schools were identified for CSI.**

² California uses a dashboard accountability system in which schools receive colors on up to seven performance indicators. Colors are determined using a combination of status (performance level) and change (the difference in school-level performance from the prior year). California's accountability system is described in more detail in the Background and Policy Context section.

Background and Policy Context

For the past quarter century, federal law has required states to measure school performance and identify the lowest performing schools. Such systems have invited critiques—that they overemphasize student achievement in reading and mathematics, exclude nonacademic features of school performance, and rely on measures closely associated with student characteristics. In December 2015, the Elementary and Secondary Education Act was reauthorized as the ESSA, ushering in a new era of education accountability.

ESSA introduced a system of multiple measures collectively intended to provide a more comprehensive picture of a school’s overall performance. Like previous policies, ESSA requires state accountability systems to include student achievement in reading and mathematics, but it expands the set of required indicators to include another academic indicator, graduation rates, progress in English language proficiency (for English learners), and a measure of school quality and student success (ESSA, 2015). ESSA also provides states with substantial latitude in determining which specific measures to include under the accountability indicators for determining school performance and how to aggregate those measures to determine which schools are underperforming.

Years after the initial implementation of ESSA accountability systems, questions remain about how well these multiple-measures systems work and, in particular, about how design choices can influence which schools are identified for CSI. Our team examined these other accountability-related questions in a larger study funded by a grant from the National Center for Education Research at the U.S. Department of Education; in this report, we focus on only one state, California.

Specifically, we examined how certain design choices regarding the measures included in California’s ESSA accountability system and the method of aggregating measures affect which schools end up in the lowest 5% and are hence identified for CSI. Drawing on administrative data from the 2017–18 school year, we conducted several analyses to better understand how California’s accountability system operates in practice, including descriptive analyses examining how many measures each school was rated on and which measures were most often missing, correlations among measures and student demographics, year-to-year correlations of measures, and simulations designed to identify the influence of specific measures on CSI identification.

Overview of California’s Accountability System

To evaluate school performance, California uses a “business rules” (or dashboard) system that assigns ratings to schools on seven possible indicators.³ For each of these indicators, schools receive a color-coded rating based on their overall status as well as their progress relative to the previous year. Schools ranked in the bottom 5% on the basis of their color-coded ratings are designated for CSI.

Describing school performance through a dashboard is intended to provide simple, easy-to-understand information. However, states must make numerous technical decisions to operationalize the calculation of the index. Each decision influences the score that schools receive and which schools are identified as low performing. These decisions include the selection of measures, the weights each measure receives, and how they are combined to determine which schools are identified for CSI.

Exhibit 1. Every Student Succeeds Act (ESSA) Indicator Performance Categories and California’s School Performance Indicators

ESSA Indicator Performance Categories	California’s School Performance Indicators
Student achievement in reading and math	Performance in English language arts and math
State-selected academic indicator for elementary and middle schools	Chronic absenteeism
Graduation rate	Graduation rate
English language proficiency progress for English learners	English learner progress
School quality or student success	Suspension rate and college and career readiness

Exhibit 1 shows how California’s selected indicators of school performance map onto ESSA’s five required accountability indicators. Because certain indicators apply to specific school levels, few schools in California would be held accountable for all seven indicators: Typical elementary and middle schools could receive ratings on a maximum of five indicators, whereas high schools could plausibly be rated on six indicators. Moreover, a school would not be held accountable for a given indicator if the data for the indicator were available for fewer than 30 students in either the current or prior year.

³ For more details regarding the underlying data and how the indicators are measured, see the *2019 California School Dashboard Technical Guide* (California Department of Education, 2019).

For each indicator, there are five possible ratings distinguished by different colors, with red indicating lowest performing and blue indicating highest performing. The color rating for each indicator is determined on the basis of two measures: the level of performance (termed *status*) and the year-to-year change in performance. A series of cut scores assigns each school to one of five categories for both performance status (how high or how low) and change in performance. Schools performing lower in both status and change would receive a lower color rating for the indicator. Exhibit 2 illustrates how a school’s ELA performance is rated according to performance status and year-to-year change. A school’s CSI status is determined by the combination of color ratings across all performance indicators. Rating combinations that result in CSI designation are as follows: red ratings for all indicators, red ratings for all but one indicator, red or orange ratings for all indicators, and five or more indicators where there majority have red ratings.⁴ In addition to being identified on the basis of performance, high schools can be identified for CSI if their two-year average combined 4- and 5-year graduation rate is below 68%.

Indicators vs. Measures

For this report, we use the terms **indicators** and **measures** to describe the elements on which schools are rated.

We use the term **indicator** to mean the main data element (e.g., math, ELA, suspensions).

We use the term **measure** to mean the specific pieces of data that contribute to overall performance on a given indicator. In the case of California, each indicator has two measures:

Status represents the level of a school’s performance on a given indicator.

Change represents how performance in the most recent year compares with performance in the prior year.

⁴ For the 2019–20 school year, ELP was included only as a status measure because year-to-year change in ELP could not yet be calculated given that a new English language proficiency assessment for English learners was implemented in the 2018–19 school year. Because change was not assessed for ELP, color ratings were not assigned. ELP counted only in a school’s combination of indicators if the status measure was “very low.” In that case, it was counted as a proxy for a red indicator. For all other statuses of ELP, it was not included at all. Because California did not assign color ratings based on ELP in the 2019–20 school year, that indicator is excluded from most of the analyses presented in this report.

Exhibit 2. Definitions of Color Ratings for the English Language Arts (ELA) Performance Indicator

Performance Level	Declined Significantly from Prior Year (by more than 15 points)	Declined from Prior Year (by 3 to 15 points)	Maintained from Prior Year (declined by less than 3 points or increased by less than 3 points)	Increased from Prior Year (by 3 to less than 15 points)	Increased Significantly from Prior Year (by 15 points or more)
Very High 45 points or higher in Current Year	Green	Green	Blue	Blue	Blue
High 10 to 44.9 points in Current Year	Green	Green	Green	Green	Blue
Medium -5 points to +9.9 points in Current Year	Yellow	Yellow	Yellow	Green	Green
Low -5.1 to -70 points in Current Year	Orange	Orange	Orange	Yellow	Yellow
Very Low -70.1 points or lower in Current Year	Red	Red	Red	Orange	Orange

Note. This exhibit is reproduced from the *2019 California School Dashboard Technical Guide* (California Department of Education, 2019, Table E, p. 189). Row headings indicate performance status, and column headings indicate performance change.

Situating California’s Approach Within the Broader Landscape of ESSA Accountability Systems

From a national perspective, California’s system for identifying CSI schools is unique in several ways. First, California is one of just a few states that uses a business-rules system to identify CSI schools. By contrast, most other states use an index-based system in which multiple measures are aggregated into a single index or score. Under this approach, states place all performance indicators on a common scale and average them, often using weights to emphasize indicators

that are deemed to have greater importance. After an overall performance index is determined for each school, those with an index value below a given threshold are designated as CSI schools.

Second, California's approach to measuring growth by incorporating year-to-year change at the school level across *all* performance indicators is distinctive. Most states use techniques that focus on growth at the student level through the estimation of student growth percentiles or value-added scores. However, those states often measure growth using only English and math assessment scores, whereas California measures growth for each of the seven performance indicators.

Third, California is one of a few states that has incorporated a measure of student discipline in its accountability system. Specifically, California has included suspension rate as a measure of school quality and student success. In regard to other indicators selected for accountability purposes, however, California is more typical. For instance, like many states, California has incorporated measures of college and career readiness and of chronic absenteeism as school performance indicators in its accountability system under ESSA.

Objectives and Research Questions

This report is part of a broader study examining the underlying theory of action of accountability in the context of ESSA. Other study components examine how principals in CSI schools approach school improvement, what support is provided to CSI schools, and whether student outcomes improve as a result of schools being identified for CSI.⁵ Our primary objective in conducting the analyses presented in this report is to better understand the choices California made when designing its ESSA accountability system and the implications of these choices for which schools are identified for CSI. In particular, we addressed the following research questions (RQs):

RQ 1: Which indicators, as defined and measured within California's accountability system, are most often used to evaluate school performance in practice? How likely are CSI schools to receive low-performance color ratings on each indicator?

Under ESSA, schools are rated on a variety of indicators, providing a more holistic measure of school performance compared with prior accountability policies, which were largely based on students' test performance. However, not all schools receive ratings on all indicators. In particular, some indicators may not apply to all schools because they are grade specific (e.g., graduation rates), and schools receive a rating only if they have sufficient numbers of students

⁵ All publications from this study are posted to the following project website: <https://www.air.org/project/impact-csi-designation-multiple-measure-essa-accountability-systems>.

contributing to the calculation of performance for the given indicator. For this RQ, we investigated the number of indicators for which schools received ratings, which indicators were most often not rated, and which indicators most often resulted in low-performance color ratings.

RQ 2: How are school performance measures related to each other and to school demographics?

ESSA requires states to include multiple indicators of school performance under the assumption that each will provide complementary—not duplicative—information. Highly correlated measures provide limited information about a school’s performance; novel measures may yield new insights and influence a school’s accountability rating. The inclusion of each additional accountability measure will have less influence on a school’s CSI status if the additional indicator is highly correlated with other indicators already included in rating school performance (Harris & Liu, 2018). In other words, measures should have more influence if they add unique information about school performance not already represented in the other accountability measures. For this RQ, we conducted an exploratory analysis to determine how distinctive each of California’s accountability indicators is from the others.

School accountability systems are intended to measure school performance, not simply to reflect features of the student population. Measures that are strongly correlated with demographics may not be good measures of school performance (Di Carlo, 2019) and may be seen as unfair—particularly to high-poverty schools (Wright & Petrilli, 2017). However, these measures may be useful for identifying schools needing high levels of support. Therefore, we also investigated the extent to which each accountability indicator was related to school demographics.

RQ 3: To what extent are ratings on measures and indicators consistent from year to year?

Measurement of school performance should exhibit some consistency over time; large fluctuations from year to year could signal noise or randomness in the measure. In such cases, the information provided by the measure may not be meaningful or reliable. Because accountability systems are intended to foster school improvement over time, one would not want perfect consistency in school performance—that would mean schools would have little hope of improving their performance ratings. However, on balance, we would expect some amount of stability over time given that academic performance tends to change incrementally. For this RQ, we examined whether the measures included in California’s accountability system exhibited consistency over time, signifying their stability.

RQ 4: What is the influence of individual performance indicators on schools’ CSI designation under California’s accountability system?

Multiple measures are included in accountability systems to provide a multifaceted perspective on school performance. If certain measures have little or no influence on schools’ ratings,

however, policymakers and educators could reasonably question whether the school ratings are indeed multidimensional. Several of the prior RQs investigated aspects of influence. Specifically, RQ 1 examined which indicators were most often unrated. All things being equal, indicators that are more often rated for schools will have more influence. In addition, RQ 2 examined the uniqueness of indicators. As explained, indicators that are more distinctive should have more influence on school ratings because they provide new and different information. For RQ 4, we examined influence more explicitly by analyzing how the exclusion of individual indicators from the accountability system would change school performance rankings and the set of CSI schools.

Method

To address the RQs about California’s school accountability system and its effects on the identification of CSI schools, we analyzed data from extant sources using a variety of analytic methods.

Data and Sample

Analyses in this report rely on school-level information on the indicators and measures used in defining school performance under California’s accountability system (e.g., test scores, chronic absenteeism, and suspension rates) obtained from the California Department of Education’s School Dashboard. We used data from the 2019 School Dashboard, which includes performance data for the 2018–19 school year (along with 2017–18 to measure change) to assign CSI designations for the 2019–20 school year. We paired accountability information with data on school demographic composition (e.g., race/ethnicity of the student population, subsidized meal eligibility, students with disabilities), total enrollment, and school level (elementary, middle, and high school) for the 2018–19 school year from the California Department of Education.

The data included all public schools in California. However, the schools included in specific analyses varied depending on the RQs addressed. Most analyses in this report included all public schools in California, overall or by school level. Analyses related to RQ 1 examined all public schools as well as CSI schools only.

Analytic Approach

Descriptive Analyses

We used descriptive analyses to investigate which accountability indicators are used to evaluate school performance (RQ 1). Specifically, we examined the number of indicators on which schools were rated and, for each indicator, the percentage of schools that received a

given color rating or that were not rated because they had an insufficient number of students contributing to the calculation of performance.⁶ Results for these analyses compare all public schools with CSI schools by school level.

We also conducted descriptive analyses to explore the connection between economic disadvantage and indicator performance (RQ 2) by examining differences in performance between schools in the highest and lowest poverty quartiles, as defined by the percentage of economically disadvantaged students. For the most common indicators, we examined differences between the two groups of schools in terms of the distribution of performance on both status and change as well as the colors assigned on those indicators for low- and high-poverty schools.

Correlational Analyses

We used simple correlational techniques to examine the relationship among performance on different indicators and between indicator performance and school demographics (RQ 2) and to investigate year-to-year consistency in indicator ratings in terms of colors assigned as well as status and change ratings (RQ 3). In addition, we performed multiple regression analyses to determine whether performance on other indicators, student demographics, and school characteristics were good predictors of indicator performance in terms of both status and change (RQ 3).

Simulation Modeling

To examine the influence of individual accountability indicators on the likelihood of CSI designation (RQ 4), we used simulations comparing the actual set of CSI schools (under existing accountability rules) with the set that would be identified if a specific indicator were excluded from school performance ratings. The difference between the set of actual and simulated CSI schools—measured as both the percentage of newly identified schools (those identified under the simulation but not using actual calculations) and the percentage of schools no longer identified (those identified using actual calculations but not under the simulation)—reflects the degree of influence for a given indicator.

In addition, we conducted a set of simulation analyses to examine whether some CSI schools could have avoided the CSI designation had they performed better on a given indicator. These results speak to which measure each school might prioritize to avoid CSI designation, assuming that performance on the other indicators remains unchanged. For example, if an improvement from a red rating to an orange rating on chronic absenteeism would lead to a school not being identified for CSI, prioritizing a reduction in chronic absenteeism would be prudent for that

⁶ For most indicators, a rating is calculated if a school has at least 30 students with reportable data for 2 consecutive years.

school. In this set of analyses, we examined for each indicator which color rating would result in a CSI school avoiding CSI designation given the color rating the school actually received. Schools would increase their chances of avoiding CSI designation by improving red indicators to orange or yellow or improving orange indicators to yellow because CSI designation is the result of some combination of red and orange indicators. Any improvement of indicators that are already yellow or better would not help with avoiding CSI designation. Likewise, if a measure is not rated, improvement on that measure is moot as it relates to CSI designation.

Results

In this section, we present the results of our analyses organized by RQ.

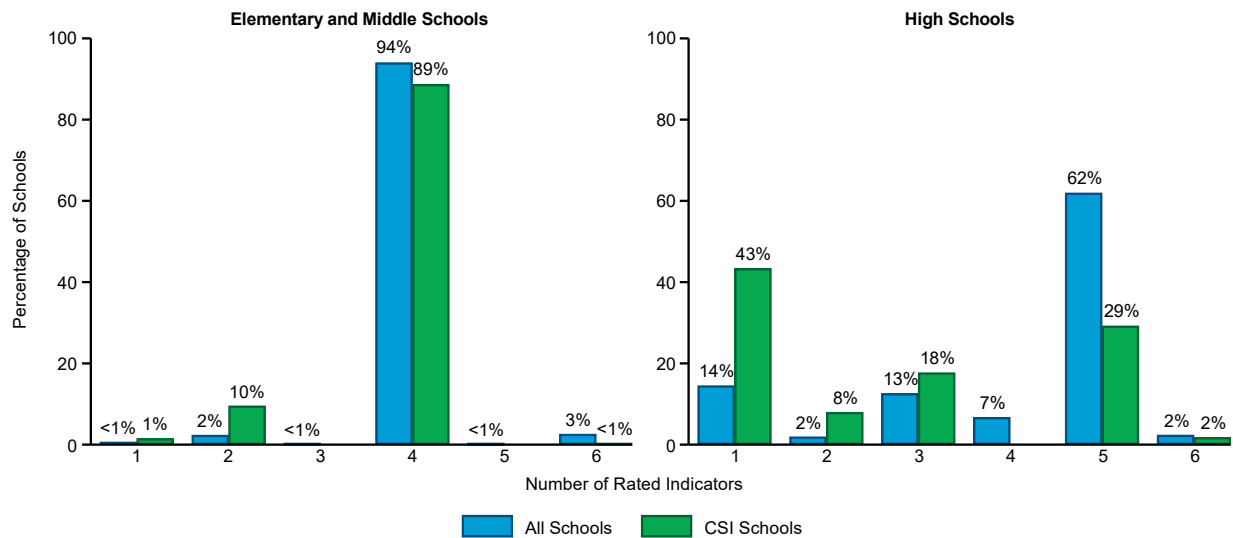
RQ 1: Indicators Evaluating School Performance in Practice

Under California's school accountability system, each school is rated on up to seven performance indicators. Given that some indicators apply only to certain grade levels, most schools are rated on four or five indicators rather than all seven possible indicators. Here, we present the result of analyses investigating the number of indicators on which schools are rated, the indicators on which schools are most often rated, and the indicators on which schools most often receive colors denoting low performance.

CSI schools often had fewer rated indicators than was typical of all public schools in California. For example, whereas only 2% of all elementary and middle schools were rated on only two indicators, 10% of CSI elementary and middle schools were rated on two indicators (Exhibit 3). This finding was particularly acute at the high school level: 43% of CSI high schools were identified on the basis of only a single indicator in 2019–20, compared with 14% of all high schools. High schools rated on a single indicator tended to be small (an average enrollment of 60 students) and identified as alternative schools (78%).

Across all grade levels, most CSI schools received a red rating on the suspension rate indicator; at the elementary and middle school levels, most CSI schools also received a red rating on the chronic absenteeism indicator. As expected, across all indicators in 2019–20, the share of CSI schools that received a red rating was higher than the share of all schools that received a red rating, but the difference was particularly striking for the suspension and chronic absenteeism indicators (Exhibit 4). For example, 54% of CSI elementary and middle schools and 73% of CSI high schools received a red rating on the suspension indicator in the 2019–20 school year, compared with 6% and 13% of all schools at the elementary/middle and high school levels, respectively, in California.

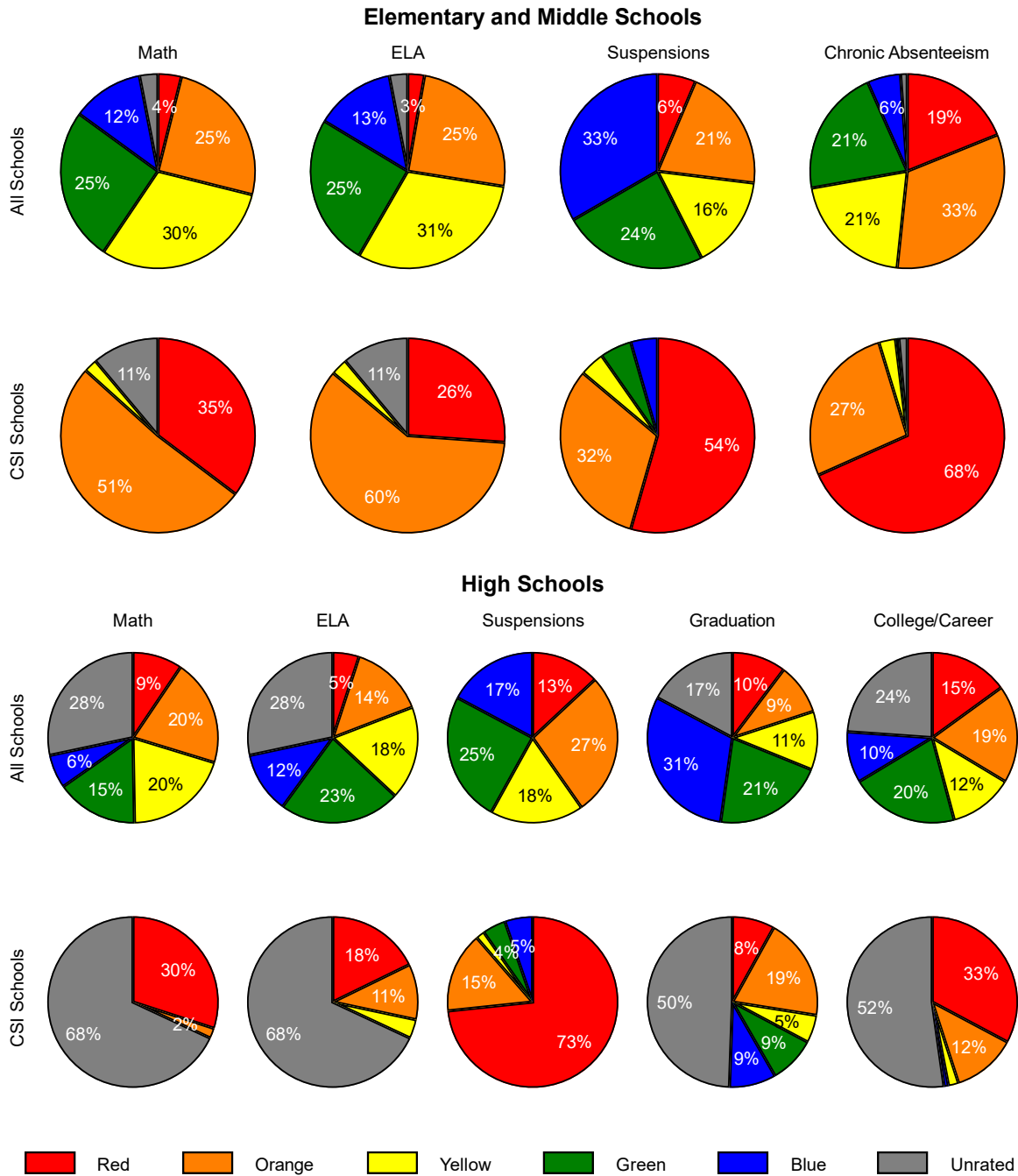
Exhibit 3. Percentage of All Public Schools and Comprehensive Support and Improvement (CSI) Schools in California by Number of Indicators Rated in 2019–20, Separately for Each Grade Level



Note. For all elementary and middle schools, $n = 7,604$; for CSI elementary and middle schools, $n = 336$; for all high schools, $n = 1,876$; and for CSI high schools, $n = 113$. The “all schools” group includes both CSI and non-CSI schools. CSI high schools include only those identified on the basis of performance on the state dashboard system rather than graduation rates below 68%. The analysis presented in this exhibit did not include English learner progress (ELP) as an indicator because schools were not assigned color ratings for ELP in 2019–20.

CSI schools were much more likely to be unrated on both ELA and math performance. In particular, because they did not meet the required number of students, 11% of CSI elementary schools and two thirds of CSI high schools were unrated on ELA and math performance, compared with 3% of all elementary schools and less than one third of all high schools (Exhibit 4). However, when rated on these achievement indicators, CSI schools were more likely to receive a red rating on math performance than on ELA performance. In addition, about half of CSI high schools identified on the basis of performance were unrated on the graduation rate and college and career readiness indicators.

Exhibit 4. Percentage of All Public Schools and Comprehensive Support and Improvement (CSI) Schools in California With a Given Color Rating for Each Indicator Rated in 2019–20, Separately for Each Grade Level



Note. For all elementary and middle schools, $n = 7,604$; for CSI elementary and middle schools, $n = 336$; for all high schools, $n = 1,876$; and for CSI high schools, $n = 113$. The “all schools” group includes both CSI and non-CSI schools. CSI high schools include only those identified on the basis of performance on the state dashboard system rather than graduation rates below 68%. The analysis presented in this exhibit did not include English learner progress (ELP) as an indicator because schools were not assigned color ratings for ELP in 2019–20. ELA = English language arts.

RQ 2: Correlations Among Performance Measures and Between Performance Measures and School Demographics

Under ESSA, the evaluation of school performance is intended to include multiple measures, moving away from the more narrow focus of the No Child Left Behind Act on mathematics and ELA proficiency. The intent is for each measure to provide different information about a school's performance, resulting in a more complete picture of school quality. Prior to ESSA, policymakers and researchers critiqued traditional measures of school performance (primarily achievement in reading and mathematics) as too closely associated with school demographics (e.g., McEachin & Polikoff, 2012), a relationship that could potentially be disrupted by the introduction of additional measures under ESSA. In this section, we describe the relationships among accountability measures and between accountability measures and school demographics.

Relationships Among Performance Measures

Correlations among status measures were much stronger than correlations among change measures. As mentioned earlier, California's accountability system considers both the status (i.e., level of performance) and the change (i.e., year-to-year change in performance) for each indicator when rating school performance. We found strong, positive correlations among the status measures of school performance indicators (see Exhibits 5 and 6). For example, the correlation between ELA performance status and math performance status was .92 at the elementary and middle school level and .94 at the high school level. Thus, these measures are nearly duplicative and provide no new information about different aspects of school performance.⁷ Furthermore, at the elementary and middle school level, all but one correlation among the status measures was at least .30. The same was true for all but two correlations among the status measures at the high school level. In contrast, most correlations among the change measures and between the change and status measures were below .30 (and often close to zero).

Among all status measures, suspension rate and ELP were the two with the weakest correlations with other measures. The correlations between the suspension status and other status measures at the high school level were $-.41$ for ELA and math performance, $-.17$ for graduation, and $-.32$ for college and career readiness.⁸ For ELP, the correlations with those status measures were of similar magnitude (.42, .30, and .36, respectively). This suggests that the suspension and ELP measures provide unique information about schools relative to other

⁷ Even though ELA and math performances statuses are nearly duplicative, including both measures independently may serve other purposes, such as encouraging schools to focus on both math and ELA performance, and may provide insight in the limited instances where performance on the two does differ.

⁸ These correlations are negative because higher rates of suspensions and chronic absenteeism are indicative of lower student outcomes, whereas higher values for math and ELA performance are indicative of higher student outcomes.

measures. In contrast, the correlations were .61 between graduation status and ELA and math statuses and .66 between graduation status and college and career readiness status at the high school level. The suspension rate and ELP status measures also had relatively low correlations with other measures at the elementary level.

Exhibit 5. Correlations Among Measures of Performance Indicators for Elementary and Middle Schools

Performance indicator	1	2	3	4	5	6	7	8
1. Math Status	—							
2. Math Change	.15	—						
3. ELA Status	.92	.09	—					
4. ELA Change	.03	.66	.08	—				
5. Chronic Absenteeism Status	-.54	-.05	-.56	-.01	—			
6. Chronic Absenteeism Change	-.10	-.05	-.13	-.07	.40	—		
7. Suspension Status	-.42	-.10	-.35	-.02	.36	-.03	—	
8. Suspension Change	.05	-.09	.05	-.10	.02	.11	.24	—
9. ELP Status	.39	.06	.43	.06	-.31	-.05	-.13	.00

Note. The number of observations differs for each correlation, ranging from 6,615 to 7,797. ELA = English language arts; ELP = English learner progress.

Exhibit 6. Correlations Among Measures of Performance Indicators for High Schools

Performance indicator	1	2	3	4	5	6	7	8	9	10
1. Math Status	—									
2. Math Change	.20	—								
3. ELA Status	.94	.19	—							
4. ELA Change	.06	.64	.16	—						
5. Graduation Status	.61	.04	.61	.02	—					
6. Graduation Change	-.08	-.07	-.06	-.01	.23	—				
7. College and Career Readiness Status	.87	.00	.84	-.09	.66	-.07	—			
8. College and Career Readiness Change	-.01	-.13	.01	-.13	.10	.22	.21	—		
9. Suspension Status	-.41	-.05	-.41	.00	-.17	-.08	-.32	-.02	—	
10. Suspension Change	.05	.00	.05	-.02	.02	-.09	.04	.00	.24	—
11. ELP Status	.42	.07	.42	.04	.30	-.07	.36	-.06	-.19	.02

Note. The number of observations differs for each correlation, ranging from 1,212 to 1,966. ELA = English language arts; ELP = English learner progress.

Suspension rate and ELP status were the most distinct indicators, as evidenced by the low percentage of variance explained by other indicators. To complement the correlational analyses, we next accounted for each indicator’s relationships with all the other indicators in the accountability system simultaneously. Consistent with the individual correlations, results showed that suspension rate (both status and change) and ELP status had the lowest amount of variance explained by other indicators (Exhibit 7). In particular, only one third and one tenth of the variance in the suspension status and change measures, respectively, were explained by the other indicators at the elementary and middle school level. The corresponding values were 27% and 8%, respectively, at the high school level. In contrast, 54% and 32% of the variance in the chronic absenteeism status and change measures, respectively, were explained by the other indicators at the elementary and middle school level, with even higher percentages of variance explained for the math and ELA performance indicators. The most unique status measure is ELP status, of which approximately 20% of the variance was explained by the other indicator measures at both the elementary/middle and high school levels.

Exhibit 7. Variance in Each Performance Indicator Explained by Measures of the Other Performance Indicators

Performance indicator	Elementary/middle schools	High schools
Math Status	.901	.924
Math Change	.492	.473
ELA Status	.896	.919
ELA Change	.484	.522
Chronic Absenteeism Status	.538	
Chronic Absenteeism Change	.317	
Suspension Status	.323	.270
Suspension Change	.100	.078
Graduation Status		.571
Graduation Change		.151
College and Career Readiness Status		.842
College and Career Readiness Change		.239
ELP Status	.193	.210
N	6,579	1,166

Note. Variance explained is the R^2 value from a regression in which the status or change on a given indicator is the outcome variable and measures of the avother indicators are the explanatory variables. ELA = English language arts; ELP = English learner progress.

Relationships Between Performance Measures and School Demographics

Status measures of school performance were moderately to strongly correlated with school demographics. Most notably, the percentage of low-income students (those eligible for free or reduced-price lunch) had a fairly strong negative correlation with test scores (–.62 for math and –.70 for ELA). Among status measures of performance indicators, suspension rate had the lowest correlations with school demographic characteristics (–.14 to .20; see Exhibit 8); the highest correlation was with free or reduced-price lunch (.20).

With few exceptions, change measures of school performance indicators were virtually uncorrelated with school demographics. Among the change measures, change in chronic absenteeism and graduation rate had the strongest correlations with school demographic characteristics, but even then the correlations were fairly weak—no more than .20 for correlations involving change in chronic absenteeism and no more than .24 for correlations

involving change in graduation rate. No other change measures had correlations exceeding .07 with any school demographic characteristic. In contrast to the status measures of achievement indicators, which were especially strongly correlated with free or reduced-price lunch, the correlations between free or reduced-price lunch and change in test scores were almost zero (.02 and .03 for math and ELA, respectively), indicating no relationship between the two.

Exhibit 8. Correlations Between School Performance Indicators and Measures of School Demographic Composition

Performance indicator	% Students with disabilities	% English learners	% Free or reduced-price lunch	% Black	% Hispanic	% Asian	% Other non-White
Math Status	-.21	-.20	-.62	-.24	-.47	.47	.23
Math Change	-.03	.05	.02	-.02	.04	.01	-.04
ELA Status	-.24	-.42	-.70	-.26	-.51	.44	.24
ELA Change	-.01	.02	.03	-.01	.02	-.01	-.02
Suspension Status	.17	-.08	.20	.16	.04	-.14	.00
Suspension Change	-.02	-.01	-.03	-.01	.00	.00	.02
Chronic Absenteeism Status	.43	.08	.37	.37	.14	-.23	-.04
Chronic Absenteeism Change	.02	.05	.13	.19	.09	-.04	-.05
Graduation Status	-.45	-.27	-.20	-.18	-.09	.21	.05
Graduation Change	-.24	-.01	.09	.03	.05	-.03	-.01
College and Career Readiness Status	-.33	-.32	-.38	-.22	-.18	.40	.09
College and Career Readiness Change	-.02	.01	.04	-.02	.04	-.05	-.02
ELP Status	-.30	-.15	-.35	-.13	-.32	.31	.17

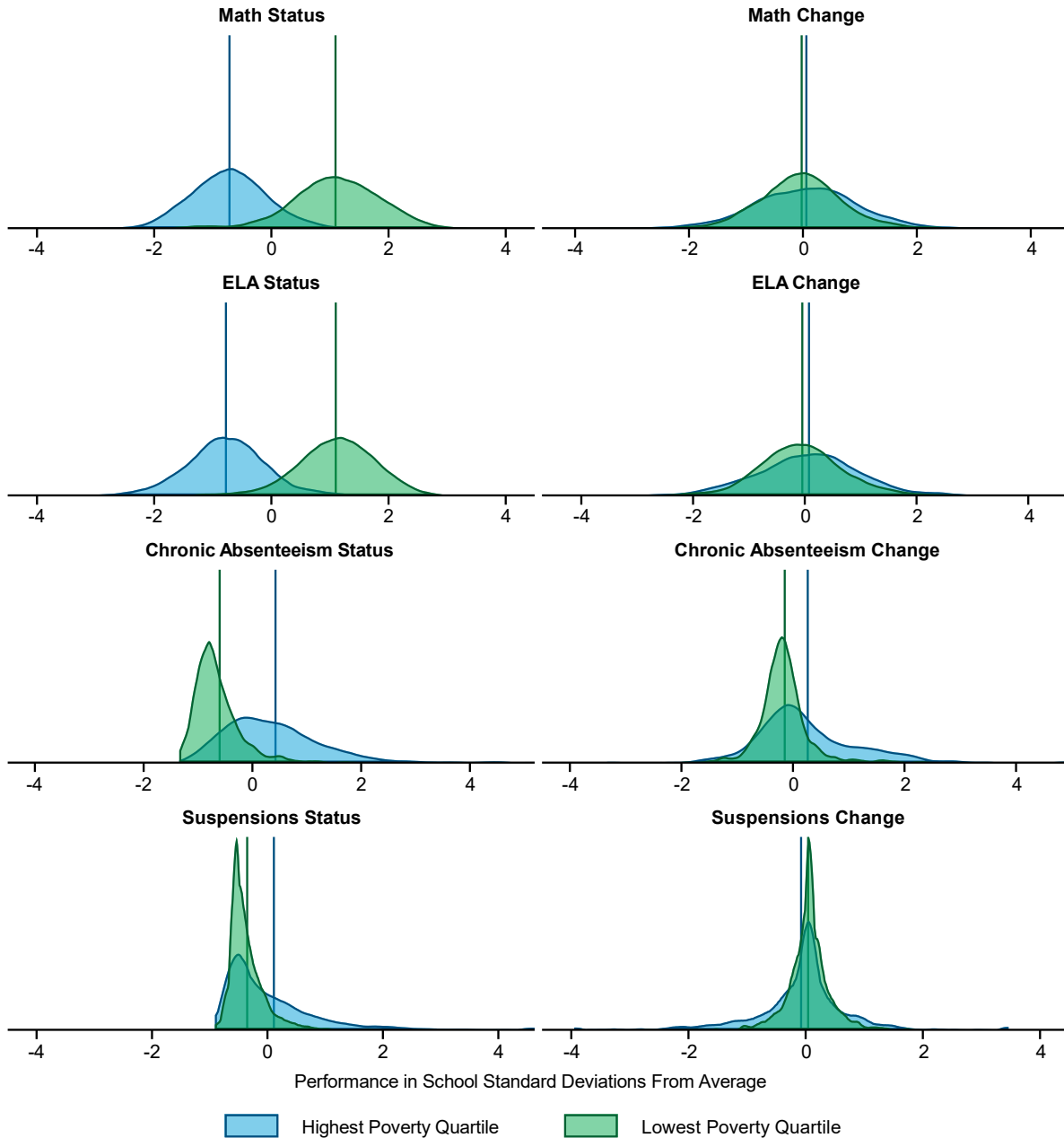
Note. The number of observations differs for each correlation: approximately 9,000 for the math, English language arts (ELA), and suspensions measures; 7,500 for the chronic absenteeism measures; 2,000 for the graduation and college and career readiness measures; and 8,000 for English learner progress (ELP).

Further analyses confirmed a strong relationship between demographics and status measures and a weak relationship between demographics and change measures. To extend the analyses of correlations, we used regression to examine the percentage of overall variance in outcomes

that is collectively explained by student demographics, along with a couple of other school characteristics (school grade level and total enrollment; see Exhibit A1 in Appendix A). Among status measures, these predictors ranged from explaining 24% of the variance for ELP to 71% for math achievement. In contrast, these predictors explained less than 2% of the variance for all change measures, except for changes in graduation rate (9.5%) and chronic absenteeism (6.5%).

On math and ELA status measures, there were practically no high-poverty schools that performed at the average of the low-poverty schools and vice versa. As depicted in Exhibit 9, for both math and ELA status measures, the average performance of high- and low-poverty schools was separated by almost 2 standard deviations, and there was very little overlap. However, for chronic absenteeism and suspension status measures, the gaps between high- and low-poverty schools were smaller. For chronic absenteeism and suspension, the averages between the two groups of schools were separated by approximately 1 and 0.5 standard deviations, respectively. In addition, for both variables, there were sizable shares of high-poverty schools that performed better than the average for the low-poverty schools, indicating a weaker relationship between student poverty and these two measures.

Exhibit 9. Distribution of Measures of Status and Change Across Schools in the Highest and Lowest Poverty Quartiles, Separately for Each Performance Indicator



Note. Vertical lines represent the average performance for the highest and lowest poverty quartiles, respectively. The highest and lowest poverty quartiles are composed of approximately 2,300 schools each. ELA = English language arts.

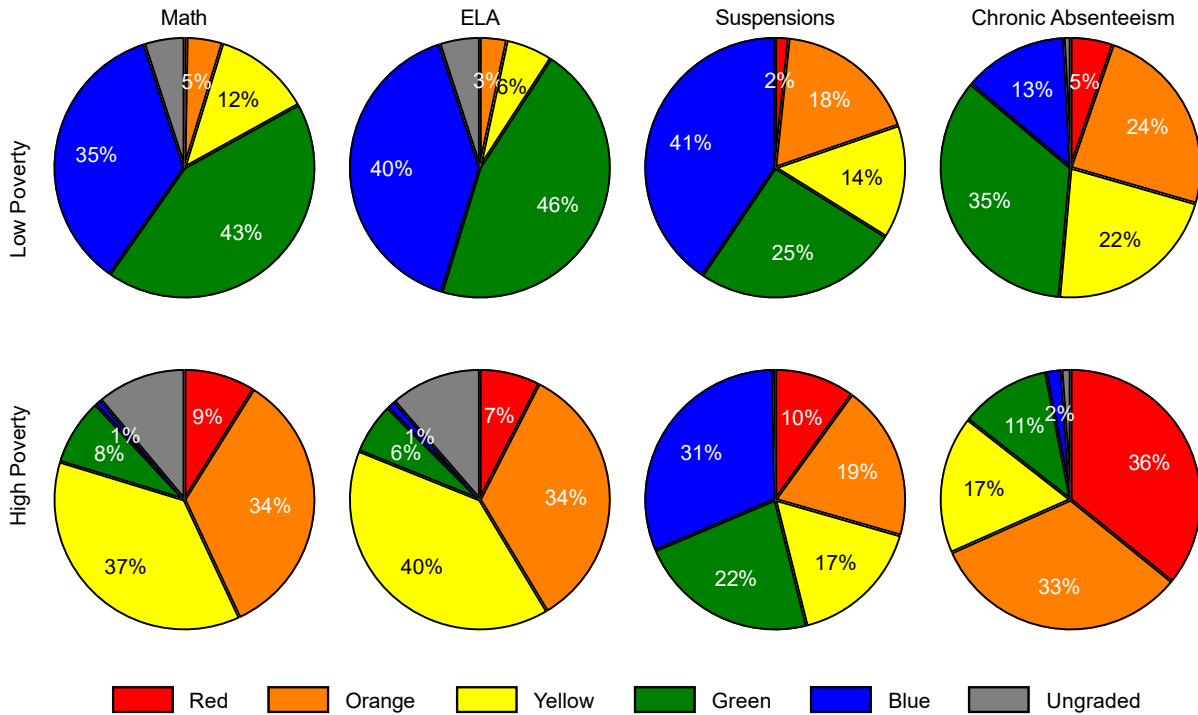
The differences between high-poverty and low-poverty schools in the performance distribution and average performance on the change measures were noticeably smaller than for status measures. For change in ELA, math, and suspensions, the averages of high-poverty and low-poverty schools were almost indistinguishable. Change in chronic absenteeism still appeared to be somewhat related to poverty, with high-poverty schools being likely to have higher increases in chronic absenteeism than low-poverty schools. But the distributions had considerable overlap, and the gap between averages of high- and low-poverty schools was much smaller compared with the gap in chronic absenteeism status.

Given the strong relationship between school poverty levels and status indicators, schools serving students from higher income families received more favorable accountability ratings. As a reminder, color ratings are the result of both status and change. Even though the inclusion of change measures might help weaken the relationships between colors assigned and demographics of schools, there are still striking differences in the colors typically assigned to low- compared with high-poverty schools. For both math and ELA, no low-poverty schools received a red rating, and only 5% and 3% received an orange rating, respectively (see Exhibit 10). Furthermore, 78% of low-poverty schools received blue or green ratings for math and 86% received blue or green ratings for ELA. In contrast over 40% of high-poverty schools received red or orange ratings for math and ELA, whereas less than 10% received blue or green ratings.

Color ratings on chronic absenteeism also differed substantially between low- and high-poverty schools. Less than 30% of low-poverty schools received a red or an orange rating for chronic absenteeism, and almost 50% received blue or green ratings. For high-poverty schools, almost 70% received red or orange ratings, and only 13% received blue or green ratings.

For suspensions, the differences by poverty level were more muted. High-poverty schools were more likely than those with low poverty to receive a red rating on suspensions and less likely to receive blue or green ratings. However, there were sizable shares of high-poverty schools that received blue or green ratings (31% received blue and 22% received green).

Exhibit 10. Color Ratings for Schools in the Highest and Lowest Poverty Quartiles, Separately for Each Performance Indicator



Note. The highest and lowest poverty quartiles are composed of approximately 2,300 schools each. ELA = English language arts.

RQ 3: Year-to-Year Consistency of Performance Indicators

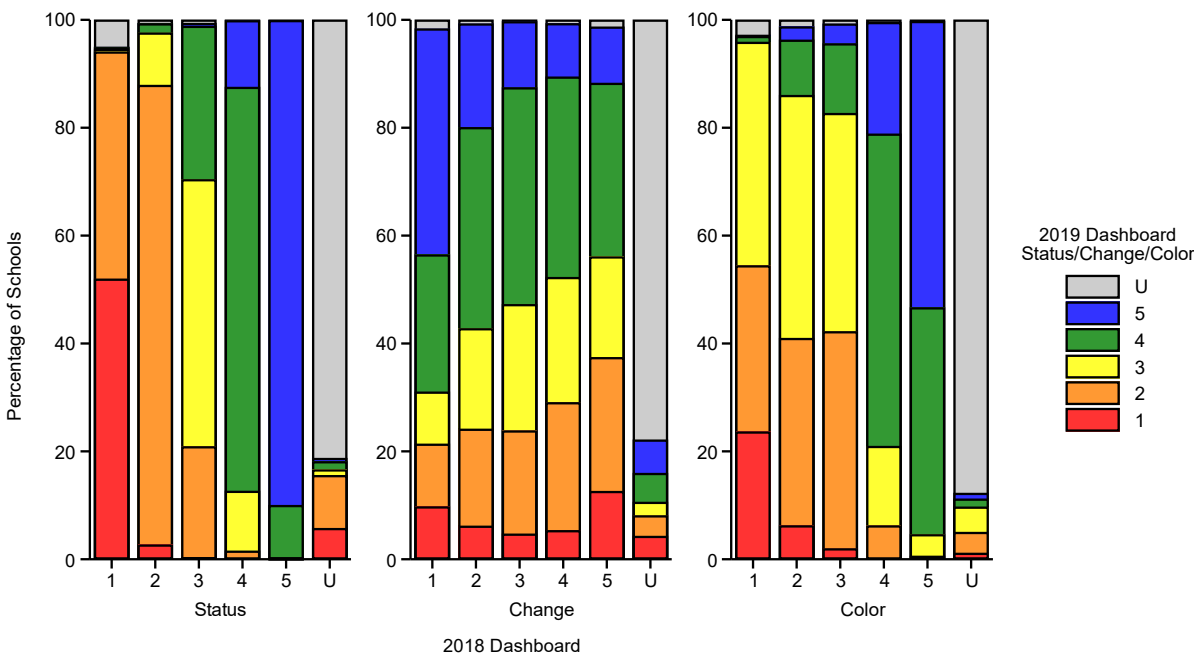
As described in the previous section, some measures—particularly the change measures—have low correlations both with other measures and with school demographics. Therefore, these measures could add unique information about school quality and could be fair insofar that all schools, regardless of the types of students served, have a chance to perform well. However, the less strong associations for these measures could be a sign that they are infused with measurement error, making the information they provide less meaningful and reliable. To better understand the degree of randomness or unreliability in these measures, we next examined the year-to-year consistency in ratings in terms of overall color as well as status and change ratings across schools.

To simplify the presentation and discussion of results, we focus on ELA performance. However, in Exhibit A2 in Appendix A, we show the analysis results for math, suspensions, and chronic absenteeism, in addition to ELA. The findings we present for ELA are consistent with those for other indicators.

The California accountability system showed strong year-to-year consistency in performance statuses. Among schools that received an ELA status rating of 1 in 2018 (see the leftmost bar of the left graph in Exhibit 11), more than 50% received a status rating of 1 in 2019 and another approximately 40 of those schools received a status rating of 2 in 2019. Negligible shares received status ratings of 3, 4, or 5, and approximately 5% of those schools were not rated on the ELA performance indicator in 2019. For each of the remaining bars in the status panel, the 2019 status rating for most schools was the same as their status rating in 2018 (e.g., a status rating of 2 in 2019 was the most likely rating for schools that received a rating of 2 in 2018), signaling strong consistency. Results across other indicators were similar, as shown in Exhibit A2.

In contrast to status measures, change ratings fluctuated from year to year. The middle graph of Exhibit 11 shows the distribution of 2019 ELA performance change ratings by 2018 ELA performance change ratings. Among schools that received an ELA change rating of 1 in 2018, less than 10% also received a 1 in 2019. In contrast, more than 40% received an ELA change rating of 5, and more than 25% received a change rating of 4. Furthermore, as 2018 ELA change ratings increased, the chances of getting a high ELA change rating in 2019 decreased. This fluctuation demonstrates inconsistency over time. It is not hard to imagine how this could happen. For schools that have a lower than typical performance in one year (marked by a low change rating), a return to typical performance levels the following year would be marked by a high change rating.

Exhibit 11. English Language Arts (ELA) Performance Status, Change, and Color Rating on the 2019 Dashboard by the ELA Performance Status, Change, and Color Rating in the 2018 Dashboard



Note. For color ratings, 1 represents red, and 5 represents blue. For performance status and change, 1 is the lowest rating, and 5 is the highest. “U” means the measure was unrated.

The incorporation of change introduced more variability in color ratings. Because color ratings are the result of both status and change, we observed some amount of consistency but less so than in statuses alone. For schools that were rated as red for ELA in 2018 (the leftmost bar in Exhibit 11), less than a quarter received a red rating again in 2019, and approximately 30% received an orange rating. The most common 2019 color rating for schools who received a red rating in 2018 was yellow in 2019 (~40%). There was relatively more consistency at the upper end of the performance spectrum. More than 50% of schools that received a green or blue rating on ELA in 2018 received a similar designation in 2019. However, color ratings from year to year were noticeably less consistent than ratings on status alone.

RQ 4: Influence of Performance Indicators

Within a multiple-measures accountability system, each performance indicator should contribute information to determine a school’s overall performance. In other words, each indicator should have some amount of influence in determining which schools ultimately become CSI schools. Even though each indicator is treated equally within California’s accountability system, some indicators may be more influential in determining school

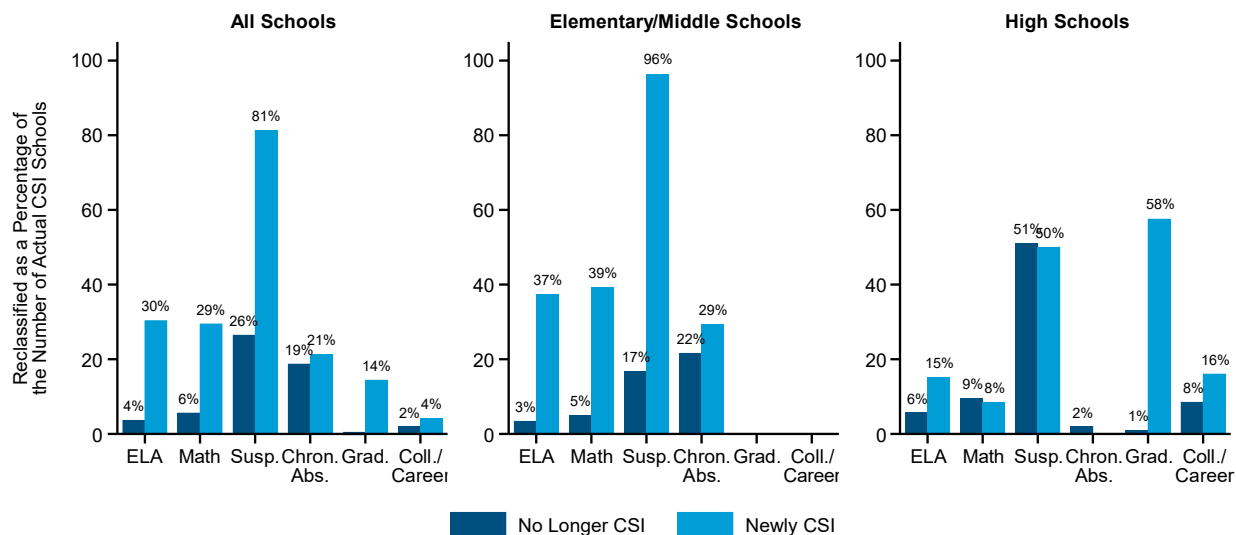
performance if they are unique or are more often rated. Here, we explicitly tested the influence of the indicators by conducting simulations that dropped individual indicators and recalculated CSI designations to see how many schools changed CSI designations under the simulation. We also conducted simulations to see how many CSI schools could have avoided their CSI designation had they performed better on a single indicator.

Simulations Dropping Individual Performance Indicators

The suspension rate indicator played a larger role than the other performance indicators in determining which schools were identified for CSI. Across all schools regardless of grade level, dropping the suspension rate indicator would result in a large number of new CSI schools (equal to more than 80% of the existing number of CSI schools; see light blue bars in Exhibit 12) and would result in approximately a quarter of existing CSI schools no longer receiving a CSI designation (see dark blue bars in Exhibit 12). In contrast, simulations dropping either the ELA or math performance indicator would result in a number of schools being newly designated for CSI equal to approximately 30% of the existing number of CSI schools and 5% of CSI schools no longer being designated for CSI, across all school levels.

These results indicate that for many schools that narrowly avoided CSI designation, their performance on the suspension rate indicator kept them from being identified for CSI. In addition, for approximately one quarter of CSI schools, suspension rate was the primary indicator responsible for their CSI designation. These patterns somewhat differed by school level. Among elementary and middle schools, the suspension rate indicator was far more likely to help schools avoid CSI designation than cause CSI designation (dropping the suspension rate indicator would result in almost as many new CSI schools as existing CSI schools). In contrast, among high schools, dropping the suspension rate indicator would result in approximately equal numbers of schools being newly designated for CSI and no longer designated for CSI—both at relatively high rates (50% of existing CSI schools).

Exhibit 12. Percentage of Schools That Would be Reclassified if a Given Performance Indicator Were Dropped From California’s Accountability System, Separately for All Schools and Each School Level

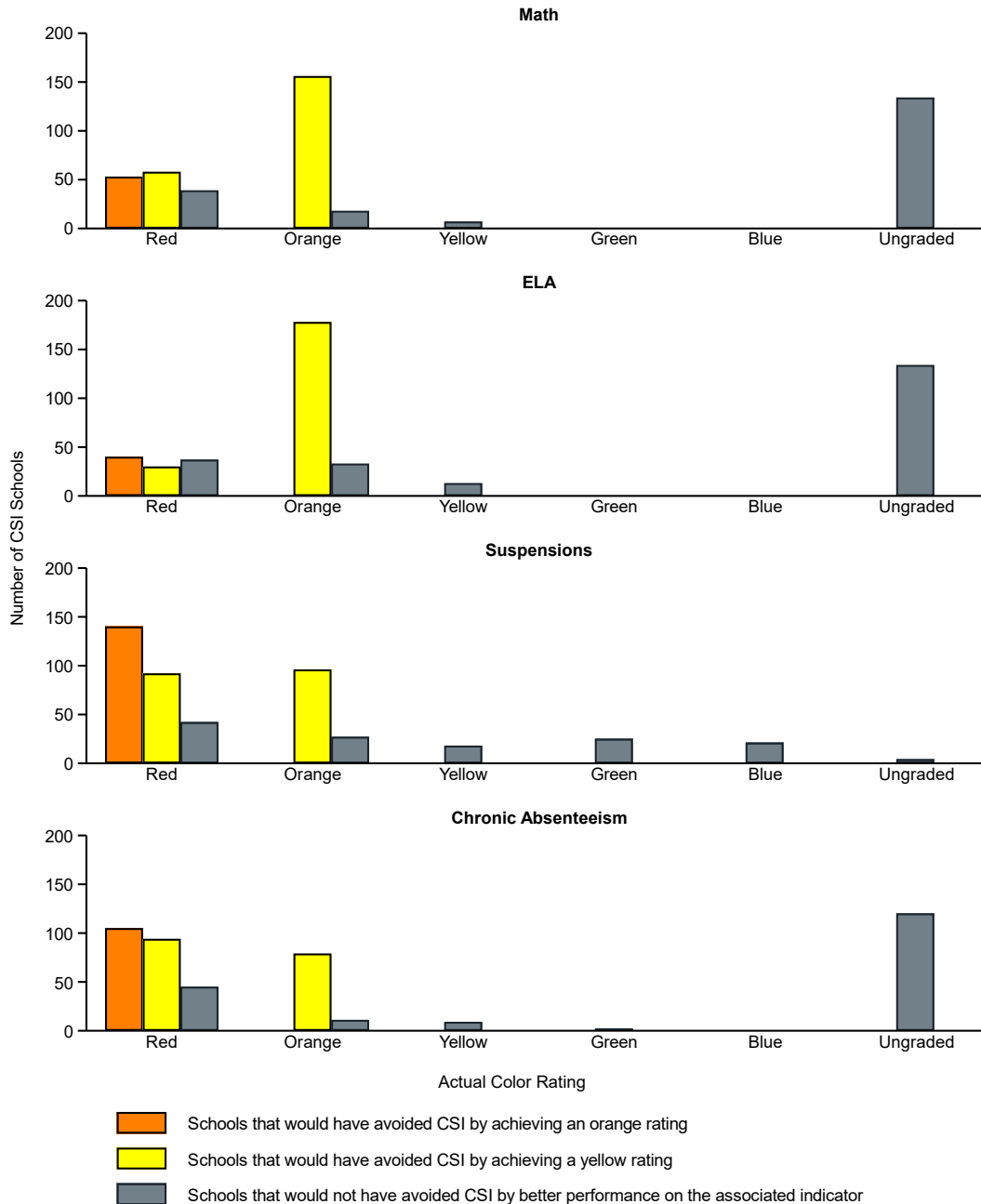


Note. Percentages were calculated using 465 comprehensive support and improvement (CSI) schools identified on the basis of performance in 2019–20 as the denominator for all schools, 324 as the denominator for elementary and middle schools, and 106 as the denominator for high schools. The values for elementary/middle schools and high schools do not account for 35 CSI schools with nontraditional grade configurations. ELA = English language arts; Susp. = suspension; Chron. Abs. = chronic absenteeism; Grad. = graduation; Coll./Career = college and career readiness.

Required Performance in Indicator Color Ratings to Have Avoided CSI Designation

Suspension rate is the indicator that CSI schools most commonly should have prioritized to have avoided CSI designation. Among the 465 schools designated for CSI on the basis of low performance (and not graduation rate), 328 would have avoided CSI designation through better performance on suspensions (by improving to orange or yellow ratings) in the 2019–20 school year (Exhibit 13). For comparison, 278 would have avoided CSI designation through better performance on chronic absenteeism. In part, this is a result of suspensions being the indicator that was least frequently unrated and most frequently given a red rating. In California’s system, red ratings carry the strongest penalty in terms of CSI designation, so focusing on achieving a high color rating on the suspension indicator (which was most often given a red rating) is a good start to avoiding or exiting CSI designation in subsequent years.

Exhibit 13. Number of Comprehensive Support and Improvement (CSI) Schools That Would Have Avoided CSI Designation With a Higher Color Rating, Separately for Each Performance Indicator



Note. N = 465 CSI schools identified on the basis of performance. This exhibit focuses on math, English language arts (ELA), suspensions, and chronic absenteeism because these are the indicators that apply to most schools identified on the basis of performance. Graduation rates and college and career readiness apply to less than one quarter of schools identified on the basis of performance. The analysis assumes performance on all other indicators is unchanged.

Discussion

When designing its accountability systems under ESSA, each state chose which measures to include and how to aggregate those measures to evaluate schools' overall performance. Those choices ultimately influenced how schools have been rated and which schools have been identified for CSI. Despite the flexibility afforded by ESSA—or perhaps because of this flexibility—the law embedded principles to anchor the design of state accountability systems. Specifically, ESSA was intended to foster state accountability systems that are comprehensive, reliable, and fair.

California designed its accountability system to be comprehensive, incorporating performance data on up to seven different indicators. To better recognize schools' performance on each indicator, leaders from the California Department of Education decided not to reduce a school's overall performance into a single overall rating. However, some of our findings suggest that this intent to measure school performance comprehensively may be compromised. In particular, CSI schools tend to be held accountable for fewer indicators than other schools—hence, their identification is based on fewer facets of school performance. Although this finding is not unique to California (Le Floch et al., 2023; Atchison et al., 2023), California's rather stringent requirement that at least 30 students in 2 consecutive years are needed to generate an indicator rating exacerbate this trend. Under California's dashboard approach to school accountability, earning a single yellow or better indicator rating substantially reduces a school's chances of receiving a CSI designation. Because larger schools will more often be rated on all applicable indicators, this means that they have more chances to earn a yellow or better rating on at least one indicator. This places small schools that are rated on fewer indicators at a disadvantage.

Our analysis also highlights the disproportionate influence of the suspension rate indicator on school identification—likely for two reasons. First, because every student counts in the measure of suspension rate and because suspension rate is measured for all school levels, almost every school receives a rating for suspensions. For small schools, suspension rate may be the only indicator or one of two indicators that the school is rated on. In these instances, a red rating on suspension rate alone places a school in CSI status. Second, suspension rate is the indicator least correlated with other performance indicators. Schools' math and ELA performance are typically closely aligned with the other indicators—and so is chronic absenteeism to a lesser extent. Strong performance on the suspension rate indicator could help otherwise low-performing schools avoid or exit CSI status, as shown by our simulations.

Some evidence exists that California's change measures are arguably more fair to high-poverty schools. Accountability under the No Child Left Behind Act was frequently criticized for reliance on student proficiency levels—which are reflective of students' poverty levels—rather than on

students' academic progress, which better reflects a school's contribution to student learning. The inclusion of year-to-year change in each performance indicator rewards schools for improvement and perhaps creates a fairer accountability system by incorporating measures that are less strongly linked to school demographic characteristics. However, the weak correlations among the change measures and between change and status measures as well as the lack of consistency in change measures over time suggest some degree of randomness and noise in these measures. To the extent that performance change is an unreliable measure of school quality, it may result in some schools that need support being overlooked and contribute to increased fluctuation in school performance ratings over time.

References

- Atchison, D., Hyland, K., Ozek, U., Le Floch, K., Hurlburt, S. (2023). *Identification of Comprehensive Support and Improvement Schools in Ohio: The Influence of Selected Measures and System Design*. American Institutes for Research.
- California Department of Education. (2019). *2019 California School Dashboard technical guide: Final version, 2019–20 school year*.
<https://www.cde.ca.gov/ta/ac/cm/documents/dashboardguide19.pdf>
- Di Carlo, M. (2019). *The false choice of growth versus proficiency*. Albert Shanker Institute.
<https://www.shankerinstitute.org/blog/false-choice-growth-versus-proficiency>
- Every Student Succeeds Act, 20 U.S.C. § 6301 (2015).
<https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf>
- Harris, D. N., & Liu, L. (2018). *What gets measured gets done: Multiple measures, value-added, and the next generation of accountability under ESSA*. Education Research Alliance.
<https://educationresearchalliancenola.org/files/publications/052518-Harris-Liu-What-Gets-Measured-Gets-Done-Technical-Paper.pdf>
- Le Floch, L., Atchison, D., Ozek, U., Hyland, K., Blair, D., & Hurlburt, S. (2023). *Multiple Measure Accountability Under ESSA: Early Findings From Three States*. American Institutes for Research.
- McEachin, A. & Polikoff, M. S. (2012). We are the 5%: Which schools would be held accountable under a proposed revision of the Elementary and Secondary Education Act? *Educational Researcher*, 41(7), 243–251.
- Wright, B. L. & Petrilli, M. J. (2017). *Rating the ratings: An analysis of the 51 ESSA accountability plans*. Fordham Institute. <https://fordhaminstitute.org/national/research/rating-ratings-analysis-51-essa-accountability-plans>

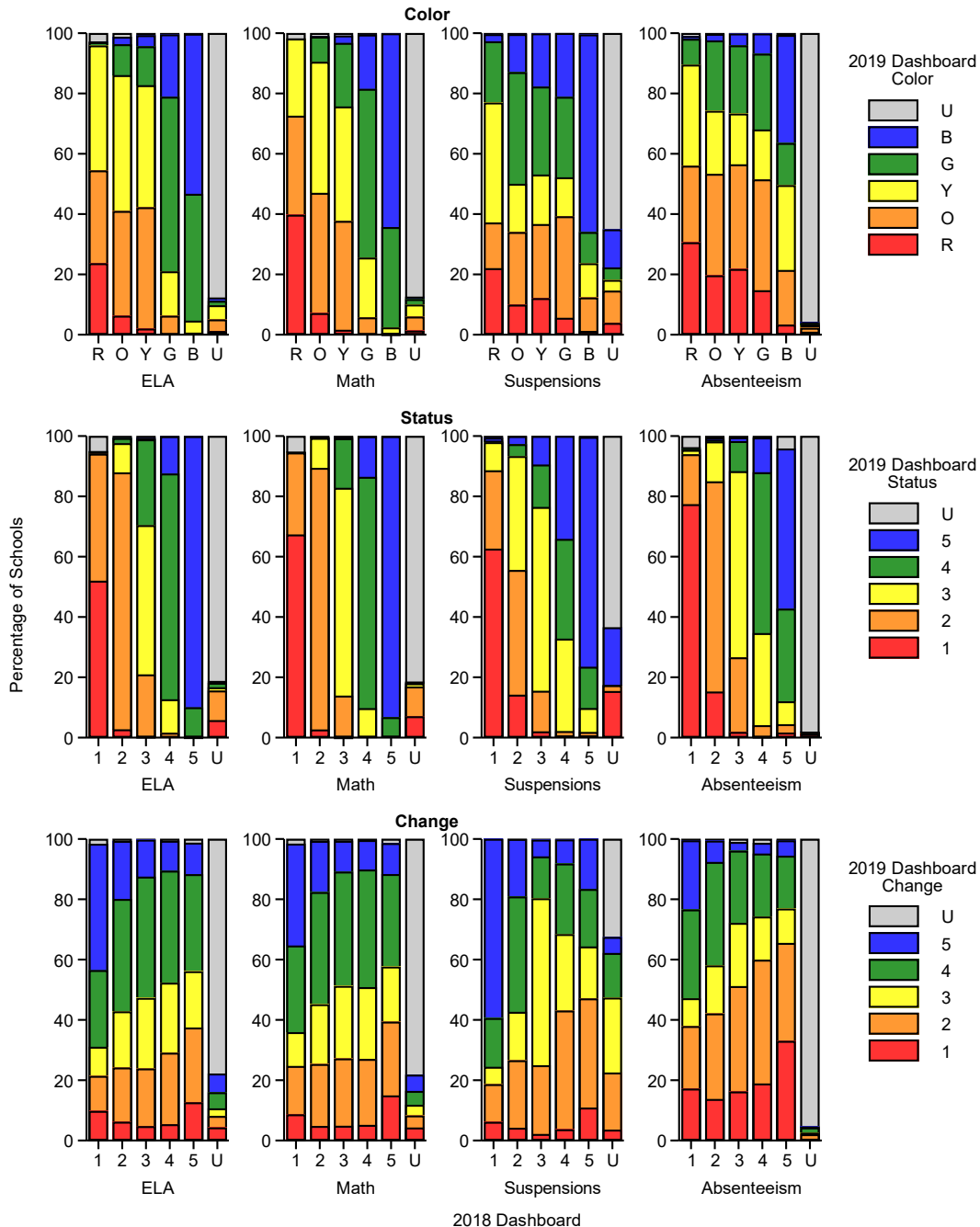
Appendix A. Supplemental Exhibits

Exhibit A1. Variance in Measures of School Performance Indicators Explained by Student and School Characteristics

Performance indicator	Variance explained (R^2)	Number of schools
Math Status	.711	7,568
Math Change	.011	7,534
ELA Status	.664	7,575
ELA Change	.005	7,540
Chronic Absenteeism Status	.484	6,394
Chronic Absenteeism Change	.065	6,383
Suspension Status	.329	7,668
Suspension Change	.010	7,668
Graduation Status	.421	1,595
Graduation Change	.095	1,562
College and Career Readiness Status	.522	1,480
College and Career Readiness Change	.017	1,448
ELP Status	.240	6,861

Note. Variance explained is the R^2 value from a regression in which the status or change on a given indicator is the outcome variable and measures of student demographics, school size, and school grade level are the predictor variables. ELA = English language arts; ELP = English learner progress.

Exhibit A2. Status, Change, and Color Rating on the 2019 Dashboard by the Status, Change, and Color Rating in the 2018 Dashboard for English Language Arts (ELA), Math, Suspensions, and Chronic Absenteeism



Note. For status and change, 1 represents the lowest rating, and 5 represents the highest rating. “U” means the measure was unrated. For colors, “R” means red, “O” means orange, “Y” means yellow, “G” means green, and “B” means blue; “U” means unrated. ELA = English language arts.

About the American Institutes for Research®

Established in 1946, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally in the areas of education, health, and the workforce. AIR's work is driven by its mission to generate and use rigorous evidence that contributes to a better, more equitable world. With headquarters in Arlington, Virginia, AIR has offices across the U.S. and abroad. For more information, visit [AIR.ORG](https://www.air.org).



AIR® Headquarters

1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | [AIR.ORG](https://www.air.org)

Notice of Trademark: "American Institutes for Research" and "AIR" are registered trademarks. All other brand, product, or company names are trademarks or registered trademarks of their respective owners.

Copyright © 2023 American Institutes for Research®. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, website display, or other electronic or mechanical methods, without the prior written permission of the American Institutes for Research. For permission requests, please use the Contact Us form on [AIR.ORG](https://www.air.org).