

## Using Surveys of Students' Social-Emotional Learning and School Climate for Accountability and Continuous Improvement

CORE-PACE RESEARCH PARTNERSHIP

March 8, 2017

Heather Hough ([hjhough@stanford.edu](mailto:hjhough@stanford.edu))  
Demetra Kalogrides ([dkalo@stanford.edu](mailto:dkalo@stanford.edu))  
Susanna Loeb ([sloeb@stanford.edu](mailto:sloeb@stanford.edu))

### Summary and Policy Implications

- Policy makers, educators, and the broader public increasingly agree that students' development of social-emotional skills is important for success in academic and life outcomes. Research provides evidence that schools can facilitate the development of these skills, both directly and through the implementation of policies and practices that improve a school's culture and climate and promote positive relationships.
- Growing confidence that schools can contribute to students' social-emotional development has led some districts and states nationwide to consider including measures of social-emotional learning (SEL) and school culture and climate (CC) in systems of school accountability and continuous improvement.
- This policy brief summarizes our recent research using data from the CORE districts—districts serving nearly one million students who have embraced systematic measurement of SEL and CC—to provide guidance for state and local policy makers about the suitability of SEL and CC surveys as school performance indicators and how they can be used in a broader set of measures.
- We find that the CORE measures of SEL and CC demonstrate validity and reliability, distinguish between schools, are related to other academic and non-academic measures, and illuminate dimensions of student achievement that go beyond traditional indicators, all initial indications of the measures' promise for informing school improvement.
- Our results also demonstrate the importance of reporting SEL and CC measures by subgroup, as African American and Hispanic/Latino students report lower SEL and CC compared to peers even within the same schools.
- While the measures of SEL and CC provide new information for school improvement, given remaining questions about the measures' sensitivity to change over time, the effect of schools on improving SEL and CC outcomes, and the potential for measures to be gamed, further research is needed to understand the advantages and disadvantages of incorporating them into higher stakes accountability systems.

Policy makers, educators, and the broader public increasingly agree that students' development of social-emotional skills is important for success in academic and life outcomes.<sup>i</sup> Research provides evidence that schools can facilitate the development of these skills, both directly<sup>ii</sup> and through the implementation of policies and practices that improve a school's culture and climate and promote positive relationships.<sup>iii</sup>

Growing confidence that schools can contribute to students' social-emotional development has led some districts and states nationwide to consider including measures of social-emotional learning (SEL) and school culture and climate (CC) in systems of school accountability and continuous improvement. Growing interest in local measurement is fueled in part by federal policy and state policy, both of which expand the range of measures that must be collected to support a more comprehensive understanding of school performance. The Every Student Succeeds Act (ESSA) of 2015 requires that states measure at least one indicator of "School Quality or Student Success," defined broadly to include measures of student engagement, educator engagement, student access to and completion of advanced coursework, post-secondary readiness, or school climate and safety. Similarly, under California's Local Control Funding Formula (LCFF) and the supporting Local Control Accountability Plan (LCAP), districts are expected to develop and report indicators representing a wide range of educational goals, and to use these measures for continuous improvement by "identify[ing] their strengths, areas where support is needed, and where support is available within the greater ecosystem of peer learning."<sup>iv</sup>

Given the interest in measuring SEL and CC for accountability and continuous improvement, there is much to learn from California's CORE districts about how such measures could be used to understand and improve school performance. The CORE districts together serve nearly a million students and are best known for the "waiver" they received from the U.S. Department of Education that freed them from some of their federal obligations under No Child Left Behind.<sup>v</sup> CORE's unique system focuses on academic outcomes alongside non-academic measures of student success including: chronic absenteeism; suspension/expulsion; students' social-emotional learning (self-management, growth mindset, self-efficacy, and social awareness); and student, parent, and school staff reports about the culture and climate in the schools (support for academic learning, sense of belonging and school connectedness, knowledge and perceived fairness of discipline rules and norms, and sense of safety).<sup>vi</sup> CORE's systematic measurement of school and student performance on SEL and CC is unparalleled and has generated widespread national interest in the field of education and in the popular press.<sup>vii</sup> In this policy brief, we summarize our recent research to provide guidance for state and local policymakers about the suitability of SEL and CC surveys as school performance indicators and how they can be used in a broader set of measures to support school improvement.<sup>viii</sup>

**Preliminary evidence suggests that SEL and CC measures are valid and reliable and can be used to distinguish schools with “below” and “above” average performance.**

While measures of school culture and climate have been in use for many years and have extensive research on their reliability and validity,<sup>ix</sup> there is less evidence on the use of SEL surveys as a school performance metric.<sup>x</sup> Thus, the CORE-PACE Research Partnership is coordinating a set of research activities designed to better understand the statistical properties of SEL measures and how they can be used alongside CC measures as a way to improve student and school performance.

Using data from CORE’s initial field test in 2014–15, researchers have established initial reliability and validity of the SEL measures, showing that (a) student-reported SEL and teacher reports on the same students are highly correlated; (b) student reports within schools are similar to reports across schools; (c) the scale reliabilities are high; and (d) reliability on the SEL scales is consistent across respondent groups.<sup>xi</sup> Together these findings provide validation that the student SEL self-reports are measuring the intended constructs and suggest that there is limited bias in the ways different student groups answer the questions.

In-depth qualitative studies can also shed light on the validity of the measures. In a study completed in the 2016–17 school year, researchers explored the implementation and use of CORE’s measurement system, assessing whether survey questions were answered honestly and administered fairly and whether educators or students were manipulating, or “gaming,” them.<sup>xii</sup> The study found broad support for SEL and CC measures and no evidence of this kind of distortive practice, though the measures were not yet widely used.

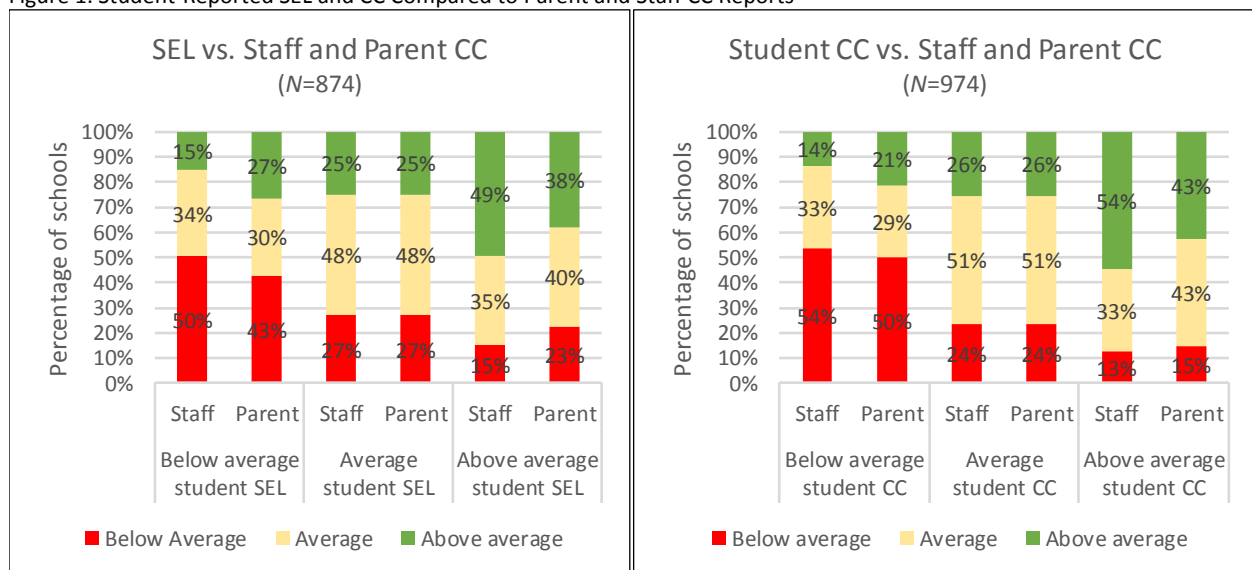
In addition to having sound measurement properties and being administered properly, in order for school-level measures to provide information about school performance, they must actually differ across schools. We find that schools do vary in both their SEL and CC reports, but that they do not vary as much in these dimensions as they do in academic achievement. As a result, while SEL and CC can distinguish the group of schools that are well above the mean from those that are well below the mean, they are not measured precisely enough to distinguish more than these blunt categories. For instance, if schools were grouped into 10 categories based on their SEL score, 84 percent of the schools in the bottom category would not be statistically different from schools in the adjacent category.

**Schools with high SEL reports also tend to have high CC reports, but schools are rarely low or high on all measures.**

Schools that have higher SEL reports also tend to have higher CC reports, not only from students but also from parents and school staff. These positive relationships are evident across elementary, middle, and high schools. This consistency in reports reinforces the idea that a positive school culture and climate contributes to students’ social-emotional learning.

However, while schools with higher SEL reports also often have higher CC reports, many schools are high on some indicators and low on others, and this variation can point to areas of improvement within schools. As an illustration, in Figure 1, we show how the parent and staff CC reports can be used with student SEL and CC reports to reveal potential problems in performance. Schools are categorized into three groups based either on student SEL (left graph) or student CC (right graph). For each of these low, middle, and high student rating groups, the graph shows the percentage of schools that are low, middle, and high for both parent CC and staff CC. For example, for the schools with a below average ranking on CC by students, 54 percent of them also ranked below average on staff CC, and 50 percent also ranked below average on parent CC. As a comparison, among schools with an above average ranking on CC by students, 13 percent are below average on staff CC, and 15 percent are below average on parent CC.

Figure 1. Student-Reported SEL and CC Compared to Parent and Staff CC Reports

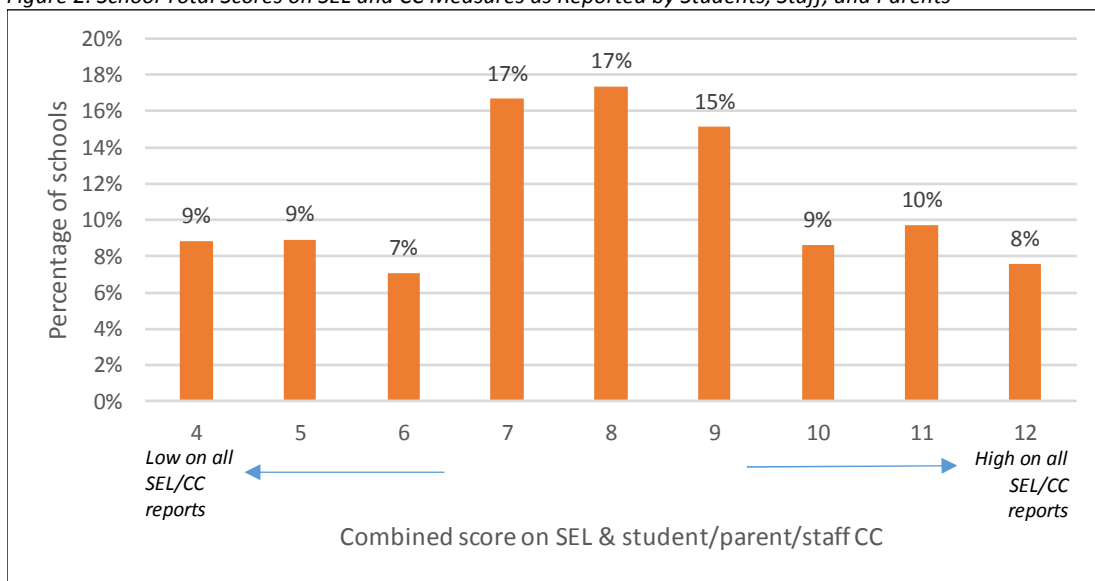


While it is easy to see that the reports by the different respondent groups tend to move in the same direction, the differences between respondent groups can be revealing when used for continuous improvement. For example, what kinds of support are needed if students feel connected to a school but their parents do not?

While differences in SEL and CC reports by respondent groups can highlight areas for improvement, when these data are used to direct precious resources to struggling schools, it may be more beneficial to use the full set of SEL and CC indicators together. For example, a district might want to identify schools that are doing poorly on all SEL and CC reports rather than just on some. To explore this idea, we added schools' scores across all of the SEL and CC indicators using a three-level scale (below average = 1, average = 2, above average = 3), such that schools that are low on all four SEL and CC indicators across respondent groups have a score of 4 and

schools that are high on all indicators have a score of 12. As shown in Figure 2, 9 percent of schools score below average (a score of 4) on all four reports; these are likely schools with a significant problem of culture/climate that may be impacting the social-emotional learning of their students. Schools with a score of 12 (above average on all indicators) may be places to investigate for evidence of excellent practice or strong leadership—ideas that can be spread to help struggling schools. Eight percent of schools fall into this category. The rest of the schools have either average scores or variation between respondents.

Figure 2. School Total Scores on SEL and CC Measures as Reported by Students, Staff, and Parents



Note. N = 1,030.

### SEL and CC reveal dramatic subgroup gaps within schools.

New measures present new opportunities to understand how schools are serving diverse students and can prompt educators and stakeholders to have honest conversations about how to develop inclusive, equitable school environments. It is for this reason that both ESSA and LCFF require disaggregation of results by subgroup. On the SEL and CC measures, we find that there are significant gaps between student groups even within schools, highlighting the need for schools to understand these disparities and work to eliminate them.

We find that students in special education, African American students, and Hispanic/Latino students report the lowest levels of SEL, and that differences between these groups persist even within schools. Figure 3 shows the gaps overall and within schools, where a score of zero indicates no difference between groups. For example, Hispanic/Latino students report an SEL score that is 0.36 standard deviations lower than white students even after controlling for other demographic characteristics. Comparing students within the same school the gaps are smaller, but still substantial (0.24 standard deviations lower than white peers in the same school).

Figure 3. SEL Gaps by Student Demographics, Overall vs. Within Schools

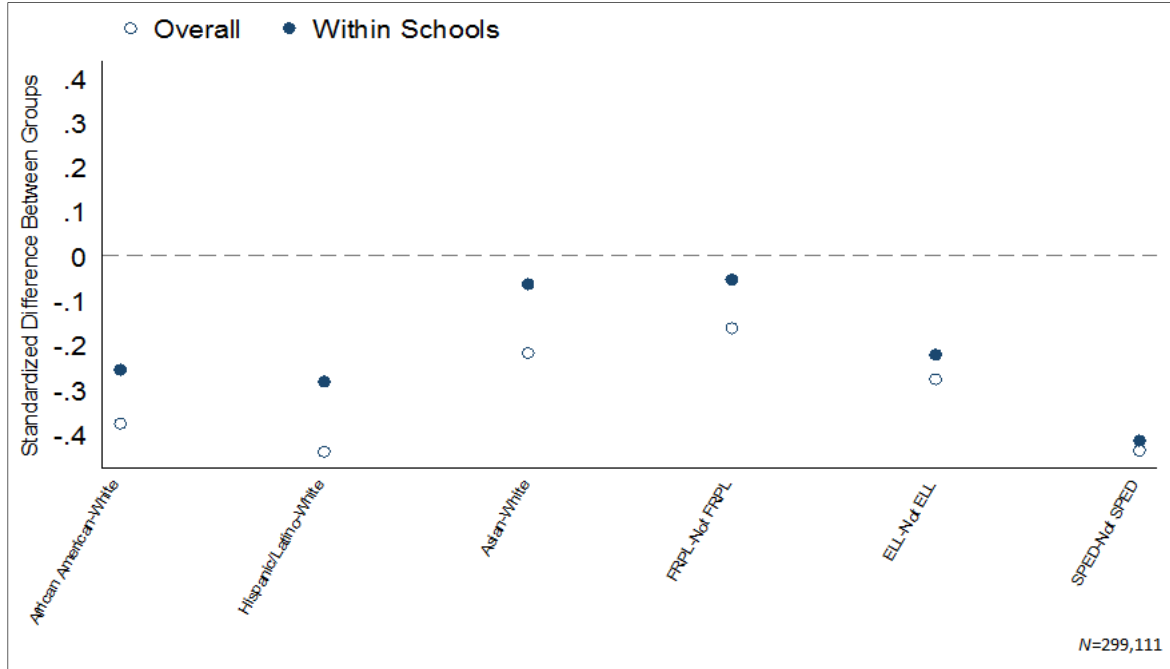
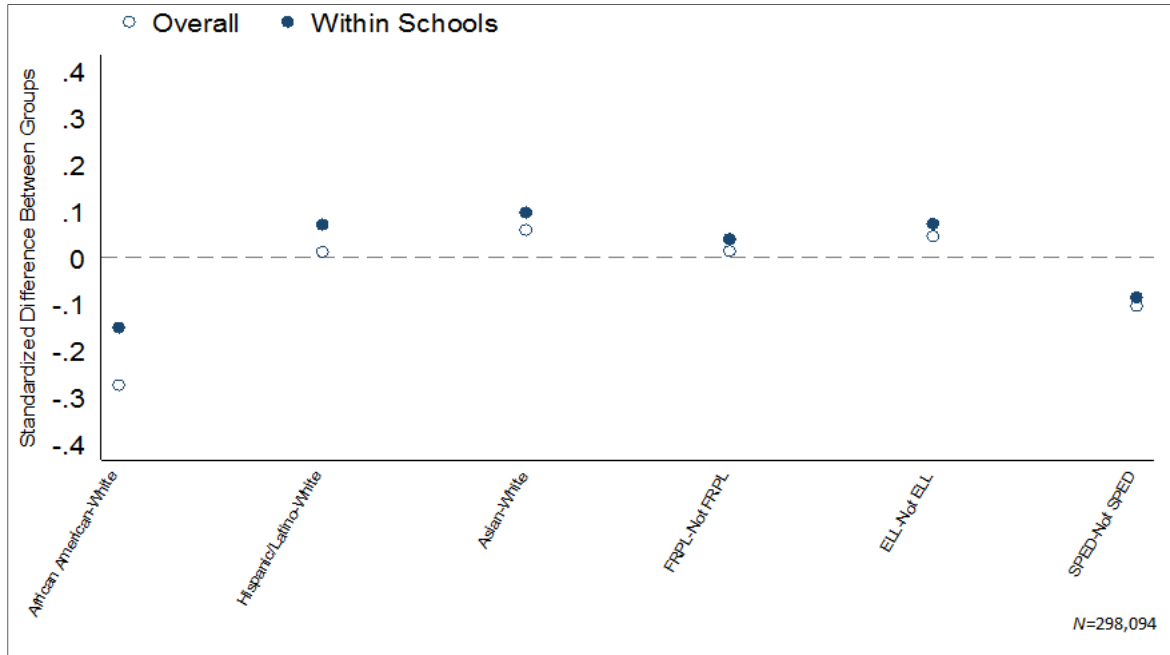


Figure 4 gives the same comparisons for CC. In this case, the only notable gap is between white and African American students, indicating that African American students report feeling less safe and supported compared to their peers *even in the same school*.

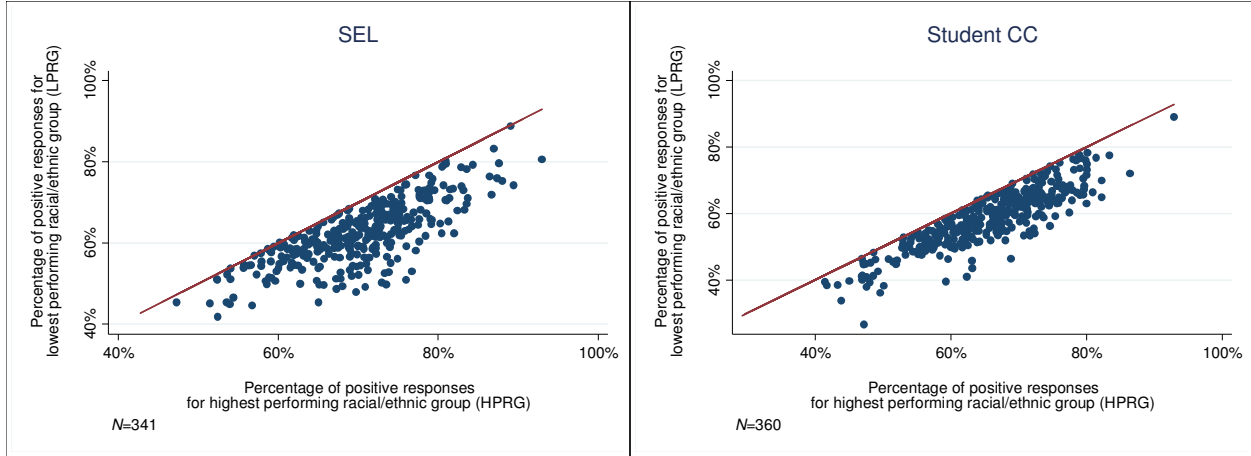
Figure 4. Student CC Gaps by Student Demographics, Overall vs. Within Schools



In addition to illuminating overall trends, the SEL and CC reports by subgroup can be used to highlight gaps within specific schools. To this end, in Figure 5, we compare the performance of the schools' lowest and highest performing racial/ethnic group (LPRG and HPRG), defined by their performance on the SEL or CC metric itself, which is a demonstration of the within-school race/ethnicity gap in SEL and CC. Each graph compares the performance of a school's highest performing racial/ethnic group (x-axis) to the lowest (y-axis), with the diagonal line demarcating schools with no gap (where the student reports for the LPRG are the same as the student reports for the HPRG), and the distance between the point for the school and the diagonal line showing the magnitude of the gap. The two indicators are highly correlated (.77 for SEL and .87 for CC), showing that school performance with the two groups is relatively similar overall. However, Figure 5 also reveals a sizable number of schools with substantial gaps between their high- and low-performing racial/ethnic groups. For example, in the school with the largest gap in SEL, 76 percent of white students (the HPRG) report positive responses, compared to only 51 percent of African American students (the LPRG). Looking at the data in this way highlights schools where improvements are needed, but also identifies schools that might be leaders in sharing promising practices. If some schools are able to close these gaps, what are they doing differently?



Figure 5. Disparities Within Schools Between Racial-Ethnic Groups in SEL and CC Reports



The subgroup gaps highlighted by new SEL and CC measures point to the need to explore why schools are seeing these disparities, and can launch a conversation about how to improve school policies and practices to better support disadvantaged students. For example, research shows that students’ experiences within school differs by race/ethnicity, including well-documented disparities in disciplinary practices and expectations for success.<sup>xiii</sup> These different experiences and treatment within schools could explain why African American students assess their schools’ culture and climate differently from their white peers, and why Hispanic/Latino students and students with disabilities report feeling less efficacious and less confident about success.

**School SEL and CC scores are predictive of academic outcomes, indicating that the measures are useful on their own and as early indicators of academic progress.**

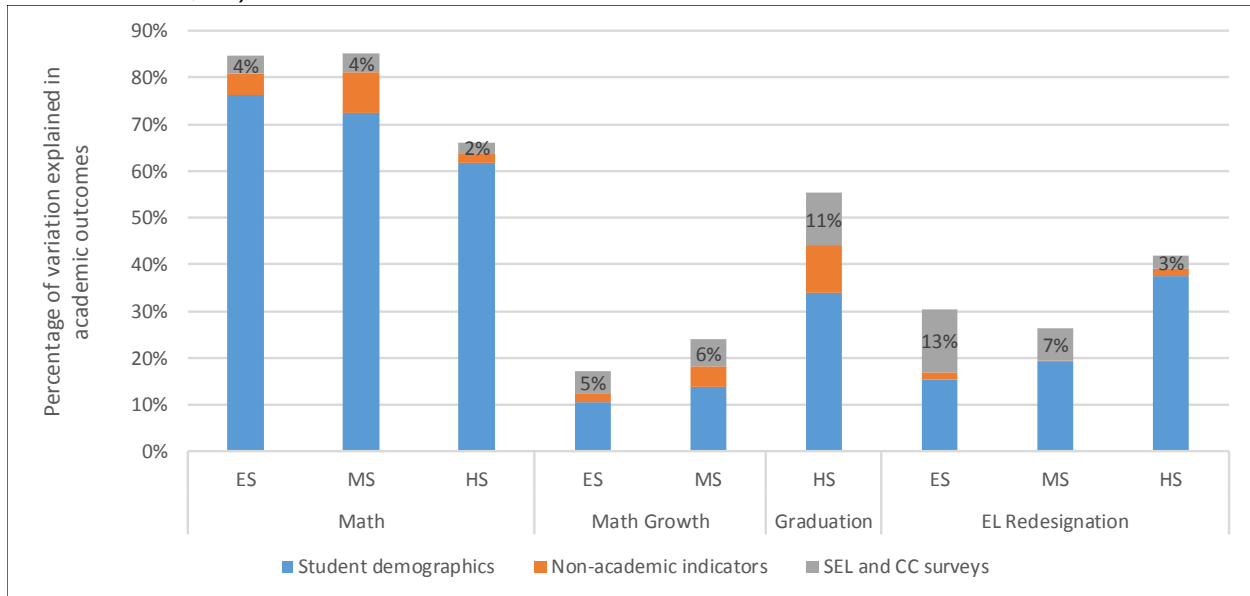
Survey-based SEL and CC measures can provide districts and states with information about non-academic outcomes in schools. While recent policy encourages the use of such measures of “School Quality and Student Success,” it also requires that the measures be related to key academic outcomes. We find that the SEL and CC measures are predictive of academic outcomes, with a particularly strong relationship to math performance on standardized tests. The SEL and CC surveys explain 54 percent of variation in math scores for elementary schools, 68 percent for middle schools, and 29 percent for high schools.

This result indicates that SEL and CC surveys are useful predictors of academic outcomes. However, because of the additional costs associated with survey administration, it is worth asking whether the survey measures provide information about schools that the administrative records do not. To explore this idea, in Figure 6, we show the proportion of variation in academic outcomes that is explained by SEL and CC surveys, after controlling for student demographics and non-academic measures derived from administrative records (chronic absence, suspension rates, and high school readiness in middle schools). We find that the SEL and CC surveys are predictive of each of the academic outcomes above available measures, although the percentage



of variation is small and variable across indicators. For example, when predicting math scores, SEL and CC survey measures combined predict 4 percent of the variation in elementary schools, 4 percent of the variation in middle schools, and 2 percent of the variation in high schools.

Figure 6. Percentage of Variation in Academic Outcomes Explained by SEL and CC Surveys, Controlling for School Demographics and Other School Quality Indicators



Note. N = 1,030.

These results show that the survey-based SEL and CC measures do indeed provide information about academic performance above and beyond the information provided by the non-academic measures available in administrative records. Additionally, the SEL and CC indicators may give insights into the processes by which schools improve their academic performance as well as how they contribute to student development in dimensions outside of the typical academic measures.

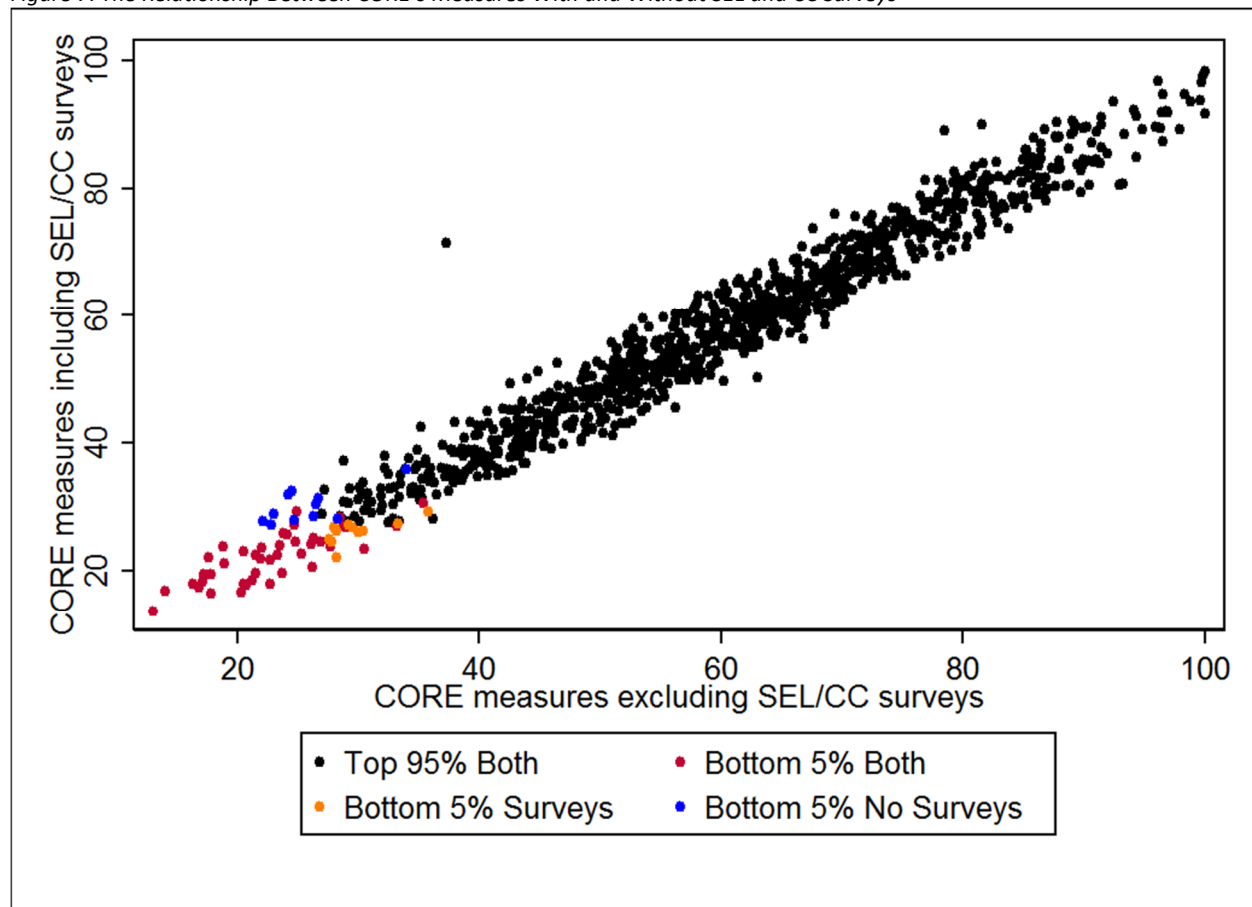
**The inclusion of SEL and CC surveys in accountability systems can change which schools are identified in the bottom 5 percent.**

One important use of school performance measures is for the identification of schools in need of support and improvement. ESSA, for example, requires that the bottom 5 percent of schools be identified for this purpose. If the SEL and CC surveys are providing important, and different, information about school performance, we should expect some change in how schools are ranked when integrating these measures into a comprehensive school measurement system.

CORE combines its multiple measures into a single summative score, and we compare how this score differently ranks schools when surveys are included compared to when they are not. The two versions of the score are very highly correlated (0.98) which indicates that, across the schools, the inclusion of SEL and CC surveys does not dramatically change the rank of schools.

However, as shown in Figure 7, the inclusion of the surveys in the summative score identifies a somewhat different set of schools in the lowest performing 5 percent. The red dots represent schools that would be identified in the bottom 5 percent of all schools with both versions, the blue dots represent schools that would only be identified without surveys, and the yellow dots represent schools that would only be identified in the bottom 5 percent of all schools once the surveys are included. We see that 53 schools are identified as being in the bottom 5 percent of all schools, but that 11 of those schools (21 percent) are different when the SEL and CC surveys are included.

Figure 7. The Relationship Between CORE's Measures With and Without SEL and CC Surveys



Note. N = 1,030.

Because SEL and CC tend to show similar trends as other non-academic and academic outcomes, school rankings based on CORE's index measure are very similar whether or not survey-based SEL and CC measures are included. While the rankings are strikingly similar, the identification of the bottom 5 percent of schools is sensitive even to small changes. As a result, the inclusion or exclusion of these survey measures affects which schools are identified in this lowest performing group. This result further points to the sensitivity of this classification and the

potential benefit of considering a wider range of measures when considering which schools are in need of improvement.<sup>xiv</sup>

**Future research is needed before new SEL and CC measures can be recommended for high-stakes use.**

In this brief, we have shown that there is good reason to pursue the measurement of SEL and CC as a way to better understand student and school performance. Measures demonstrate reliability and validity, distinguish between schools, are related to other academic and non-academic outcomes, and also illuminate aspects of student achievement that go beyond traditional indicators.

The patterns we see in the SEL and CC data suggest that the new measures can be quite powerful in illuminating a dimension of school performance that has been invisible in data systems until now. When used together in comprehensive performance measurement systems, the new SEL and CC measures can inform a deeper understanding of a school's strengths and weaknesses and prompt action on a new dimension.

However, these results are just a starting point for understanding survey-based measures of SEL and CC and how they might be used in systems of school accountability and continuous improvement. Several important dimensions of these new indicators are yet unexplored. Because the measures have not been used for high-stakes decision making, we do not know the extent to which educators will counterproductively game the measures in such a setting. As a clear example, if teachers were rewarded for positive reports of CC, they would likely give more positive reports of CC in their own survey responses. Additionally, if information revealed by these indicators will be used to prompt action, as in the identification of schools for improvement, research must first establish that the measures are sensitive to change over time and that schools themselves can and do contribute to students' growth. Perhaps even more importantly, if SEL and CC measures are to be used to support authentic continuous improvement, schools and districts must have clear guidance about what to do to improve the measured outcomes, which requires the collaboration of researchers and practitioners in schools, districts, and beyond.

## Acknowledgements

We would like to thank the many individuals who contributed to this report. First, we are grateful to the generous sponsor of this research, the S.D. Bechtel, Jr. Foundation. We also thank all of the leaders and administrators in the CORE districts for their support, feedback, and guidance throughout this project. This project would not have been completed without the assistance of many colleagues, including especially David Plank, Emily Penner, Joe Witte, and Mattias Sparrow.

The research featured in this paper is part of the CORE-PACE Research Partnership, through which Policy Analysis for California Education (PACE) has partnered with the CORE districts to conduct research designed to support them in continuous improvement while simultaneously helping to improve policy and practice in California and nationwide. Through this partnership, PACE coordinates and executes research with partners from all of California's top universities, including Stanford University, the University of Southern California, and the University of California, Davis, in addition to engaging researchers from universities and research organizations nationwide. For more information and to access all Partnership publications, visit <http://www.edpolicyinca.org/projects/core-pace-research-partnership>.

- 
- <sup>i</sup> See, for example: Nagaoka, J., Farrington, C. A., Ehrlich, S. B., & Heath, R. D. (2015). *Foundations for Young Adult Success: A Developmental Framework*. Concept Paper for Research and Practice. University of Chicago Consortium on Chicago School Research.
- <sup>ii</sup> See, for example: Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child development*, 82(1), 405–432.
- <sup>iii</sup> See, for example: McCormick, M. P., Cappella, E., O'Connor, E. E., & McClowry, S. G. (2015). Context matters for social-emotional learning: Examining variation in program impact by dimensions of school climate. *American journal of community psychology*, 56(1-2), 101–119.
- <sup>iv</sup> California Department of Education. (2016). *Preparing All Students for College, Career, Life, and Leadership in the 21st Century*. Sacramento, CA.
- <sup>v</sup> The CORE waiver districts include Fresno, Long Beach, Los Angeles, Oakland, San Francisco, and Santa Ana Unified School Districts. Sacramento City and Garden Grove Unified School Districts are also members of CORE but did not participate in the waiver.
- <sup>vi</sup> For more detail on the SEL constructs selected for inclusion, rationale for inclusion, and full survey instruments, see <http://www.transformingeducation.org/measuringmesh/>. For full CC instruments, see <http://coredistricts.org/core-index/>.
- <sup>vii</sup> See, for example: Zernike, K. (2016, February 29). Testing for joy and grit: Schools nationwide push to measure students' emotional skills. *The New York Times*. Retrieved from <http://www.nytimes.com/2016/03/01/us/testing-for-joy-and-grit-schools-nationwide-push-to-measure-students-emotional-skills.html>
- <sup>viii</sup> The full paper can be found at <http://edpolicyinca.org/publications/using-sel-and-cc>.
- <sup>ix</sup> See, for example, Berkowitz, R., Moore, H., Astor, R. A., & Benbenishty, R. (2016). A Research Synthesis of the Associations Between Socioeconomic Background, Inequality, School Climate, and Academic Achievement. *Review of Educational Research*. doi:0034654316669821.

- <sup>x</sup> For a thoughtful treatment of the concerns in measuring SEL for school accountability, see Duckworth, A., & Yeager, D. S. (2015). Measurement matters: assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237–251.
- <sup>xi</sup> West, M. R. (2016) Should non-cognitive skills be included in school accountability systems? Preliminary evidence from California's CORE districts. Vol. 29. *Evidence Speaks*. Brookings Institution. Retrieved from <https://www.brookings.edu/research/should-non-cognitive-skills-be-included-in-school-accountability-systems-preliminary-evidence-from-californias-core-districts/>
- <sup>xii</sup> Marsh, J., Bush-Mecenas, S., & Hough, H. J. (2016). *At the Forefront of the New Accountability Era: Early Implementation Findings from California's CORE Districts*. Policy Analysis for California Education. Retrieved from <http://www.edpolicyinca.org/publications/at-the-forefront>.
- <sup>xiii</sup> Bankston, C. L., & Zhou, M. (2002). Being Well vs. Doing Well: Self-Esteem and School Performance Among Immigrant and Nonimmigrant Racial and Ethnic Groups. *International Migration Review*, 36(2), 389–415.
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: two sides of the same coin? *Educational Researcher*, 39(1), 59–68.
- Lareau, A., & Horvat, E. M. (1999). Moments of social inclusion and exclusion: Race, class, and cultural capital in family–school relationships. *Sociology of Education*, 37–53.
- Lewis, A. (2003). *Race in the schoolyard: Negotiating the color line in classrooms and communities*: Rutgers University Press.
- Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A Vicious Cycle: A Social–Psychological Account of Extreme Racial Disparities in School Discipline. *Perspectives on Psychological Science*, 11(3), 381–398.
- Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99(2), 253.
- Warikoo, N., & Carter, P. (2009). Cultural explanations for racial and ethnic stratification in academic achievement: A call for a new and improved theory. *Review of Educational Research*, 79(1), 366–394.
- Watamura, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double jeopardy: Poorer social-emotional outcomes for children in the NICHD SECCYD experiencing home and child-care environments that confer risk. *Child development*, 82(1), 48–65.
- <sup>xiv</sup> In a recent analysis, Hough et al. (2016) explore the tradeoffs in different ways of identifying schools in the bottom 5 percent. In this paper, the authors show that the use of a single score is problematic because schools are rarely high or low on all measures, but rather the multiple measures illuminate strengths and weaknesses across multiple dimensions. That same caveat should be applied to the use of a single measure in this illustration. For more information, see <http://www.edpolicyinca.org/publications/identity-crisis-multiple-measures-and-identification-schools-under-essa>.