May 2019

Making Sense of Social-Emotional Survey Results Using the CORE Districts' Benchmarking Data

Katie Buckley, Ed.D.
Transforming Education

transforming education

PACE
Policy Analysis for California Education

Making Sense of Social-Emotional Survey Results Using the CORE Districts' Benchmarking Data

There has been a growing interest among districts and schools to expand their definition of student success to focus on the whole child, yet many still lack the measures needed to prioritize and inform this work. The CORE Districts, a group of districts in California committed to measuring and supporting an expanded definition of student success, have created a survey to identify the social-emotional strengths and needs for students in grades 4-12. This survey asks students about their perceptions of their own growth-mindset, self-efficacy, self-management, and social-awareness. Through Transforming Educations' resources and the Assessment Work Group Assessment Guide, the survey is freely accessible to any school or district seeking to administer it. In this paper, we provide benchmarking data for this survey, including means and standard deviations by grade-level, subgroup and competency, from nearly half a million students in grades 4 through 12 across the CORE districts, who took the survey in the 2015-16 school year. This data allows practitioners to compare their aggregated data across grade-levels and subgroups for a given scale to the appropriate CORE benchmark data, in order to make inferences about their students' social-emotional competency and mindset development. This paper concludes with three concrete ways that practitioners can use benchmark data to target resources and supports needed most within their schools and districts, based on Transforming Education's work with schools and districts throughout the country.

Acknowledgements

I am grateful to the generous sponsor of this brief, the Walton Family Foundation. I also thank all of the leaders and administrators in the CORE districts for their support and feedback. This brief would not have been completed without the assistance of my colleagues at Education Analytics and Policy Analysis for California Education.

Table of Contents

Introduction

There has been a growing interest among districts and schools in expanding their definitions of student success to focus on the whole child. This trend is supported by over two decades of research, showing that the cognitive, social, and emotional dimensions of learning are deeply intertwined. When young people develop these interconnected sets of competencies, they are more likely to be academically successful, physically healthy, engaged in their communities, financially secure, and empowered to pursue goals of their own choosing.[i]

Although schools and districts across the country are working to foster the whole child, many still lack the measures needed to prioritize and inform this work. There are a range of existing and emerging measures that can aid schools' efforts to support whole child development, including surveys of students' social-emotional competencies, school climate surveys, and chronic absenteeism data. In this brief, we focus on the CORE districts' social-emotional survey, which is one of the few validated surveys that exist to measure student perceptions of their own social-emotional competencies and mindsets.[ii] We offer rationale for how schools and districts can utilize this measure to inform practice, through the use of the CORE districts' benchmark data.

The CORE districts are a group of eight California school districts that are committed to working together to identify, implement, and scale new strategies which help students succeed. The CORE districts represent over one million students and 1,500 schools across Fresno, Garden Grove, Long Beach, Los Angeles, Oakland, Sacramento, San Francisco, and Santa Ana Unified School Districts. Beginning in the 2014-15 academic year, the CORE Districts administered a social-emotional learning (SEL) survey to all students in grades 4-12, asking students about their perceptions of their own growth-mindset, self-efficacy, self-management, and social-awareness. Within each scale, students are asked to respond to a series of items on a 5-point Likert scale (e.g., strongly disagree to strongly agree). The survey results are intended to contribute to a more holistic set of data on student development and success to inform continuous improvement in the CORE districts.

Schools in the CORE districts, as well as throughout the US, are using the CORE districts' survey to identify student social-emotional strengths and needs. Through Transforming Educations' resources and through the Assessment Work Group Assessment Guide, the survey is accessible to any school or district wishing to administer it. However, interpreting student SEL survey data can be difficult without benchmarking norms (i.e., average results by construct, grade level and student subgroup) to compare a school or district's student scores. Benchmark data from the CORE districts' survey can help educators effectively interpret survey results and take appropriate actions to foster student development.

In this paper we provide benchmarking data, including means and standard deviations by construct, grade level, and subgroup, as well as examples of how to use these data in practice. The data come from nearly half a million students across the 8 CORE districts, in grades 4 through 12, who took the survey in the 2015-16 school year. While not a true national sample, the CORE benchmarking sample is large and diverse enough to serve as a proxy for a nationally-

normed sample. As such, it can provide comparative data for other schools across the country that choose to administer the CORE districts' survey.

Development of CORE Survey

In 2013, the CORE districts applied for and received a waiver from the U.S. Department of Education which provided its member districts flexibility from key requirements of the school accountability system prescribed by the federal No Child Left Behind (NCLB) Act. Through this waiver, CORE sought to implement a new type of accountability system that, rather than looking solely at test scores and graduation rates, incorporated school performance across a broader range of outcome measures. In particular, the participating districts wanted to include measures of social-emotional (SE) skills and school culture/climate (CC), alongside the traditional academic indicators in efforts to provide a more holistic index of school quality.

After receiving their NCLB waiver in August 2013, the CORE districts sought to determine which set of SE competencies to include in their accountability system. In November 2013, CORE convened representatives from each of the participating districts, ranging from superintendents to directors of student supports, directors of social-emotional learning, and directors of special education. Content experts in this area also joined the convening from the Collaborative for Academic, Social and Emotional Learning (CASEL), the John W. Gardner Center for Youth at Stanford, and Transforming Education (TransformEd). Based on input from the convening attendees, CORE prioritized four specific SE competencies: growth mindset, self-efficacy, self-management, and social awareness.

TransformEd then curated a set of survey scales measuring these four competencies that had been developed and validated by researchers across the country. In these surveys, students in Grades 4 through 12 rate themselves across 25 questions using a 5-point Likert scale.[iii] The data from these SE surveys were combined with other academic, behavioral, and administrative data to form CORE's multi-measure indicator of school quality. In this system, academic indicators account for 60 percent of a school's score, while SE and school CC factors account for the remaining 40 percent.

The system was rolled out in stages over the course of three years, with initial SE data collection in the 2014–15 school year. Now, under Every Student Succeeds Act (ESSA), which has increased flexibility for states and districts to expand the definition of student success, CORE continues to use this measurement system to drive continuous improvement at both the district and school levels.

For more information on the curation of the SE surveys and the development of CORE's continuous improvement system, please see TransformEd's case study *Expanding the Definition of Student Success: A Case Study of the CORE Districts*[iv] and an overview of the partnership by Policy Analysis for California Education (PACE). For more information on the appropriate use of SE and CC survey data, see PACE's brief titled, *Surveys of Students' Social-Emotional Skills and School Climate for Accountability and Continuous Improvement*.[v]

Survey Validation

As the first large-scale effort in the country to systematically assess students' SE skills, the CORE Districts provide a unique opportunity to understand the properties of SE survey measures. To this end, researchers affiliated with PACE have undertaken studies that examine the quality of the survey measure in order to help clarify the ways in which students' SE skills can be measured and utilized.

In one PACE working paper, the authors offer a pragmatic approach for exploring the validity of survey-based measurement of students' SE skills.[vi] In line with prior literature on measure validation, the authors view validation as an ongoing effort involving the accumulation of evidence to ensure that a measure assesses what it is intended to measure in a particular context, for a particular population, and for a particular use. Validation evidence should be collected according to the following criteria:

- Is there evidence of **reliability**: Is the measure consistently measuring the same underlying construct?
- Does the measure demonstrate **content validity**: Is it asking the right questions in order to measure the given construct?
- Does the measure demonstrate **face validity**: Does it clearly signal the construct it is purporting to measure?
- Does the measure demonstrate **structural validity**: Do the items or tasks within the measure align with the construct(s) that they were intended to assess in the manner that is expected?
- Are the items **representative** of the entire construct: Do the items only measure a narrow section of the construct, or do they fully capture the breadth and depth of the underlying construct?
- Did survey design adhere to **best practices** in order to minimize measurement error?

The degree of evidence can vary based on the use of the survey results (i.e., measures used for high stakes require a greater degree of evidence to ensure that the measure is assessing what it is intended to assess in a reliable way). Choosing a measure by considering these six criteria can help researchers determine whether the measure is suitable for their intended purpose. There are several PACE-affiliated studies that examine the validity and reliability of the CORE SE survey along these criteria. We highlight a few key studies below.

Based on findings from the 2014-2015 field test with nearly 500,000 students, student ratings on each of the measures were reliable and correlated in the expected direction with other academic and behavioral outcomes.[vii] For example, students' self-efficacy ratings correlated positively with GPA and standardized Mathematics and ELA scores and correlated negatively with the number of days a student was suspended as well as the total number of days s/he was absent. In other words, students with stronger SE skills tend to have better grades and test scores and are less likely to be absent or suspended throughout the school year.

In another PACE working paper from researchers at Education Analytics, the authors used classical test theory and item response theory to explore the measurement properties of these surveys.[viii] While their findings suggest that the scales have reasonable measurement properties, the authors provide several recommendations to improve the surveys, including: rewording

negatively phrased items, continuing to explore whether any of the items are interpreted differentially across student subgroups, replacing items every few years, and creating developmentally appropriate items that differ across grade-levels.

In a separate PACE report, researchers explored how schools can use these measures to support students' social-emotional and academic development.[ix] The authors found that in the CORE Districts, the SE surveys are capable of distinguishing between schools with high and low self-reported student SE skills. The authors also found that student SE scores are predictive of academic outcomes and can explain the variation in these outcomes beyond the information provided by other non-academic measures available to schools. Furthermore, SE scores can highlight gaps in perceptions among subgroups of students within schools. The findings suggest that students' self-report on these measures can be useful for driving continuous improvement in and across schools.

Based on these studies and several others, researchers Hunter Gehlbach and Heather Hough conclude that "the question of whether a particular district should incorporate student perception surveys into its assessment system will depend upon a host of factors. Inevitably, smart decisions will depend on nuances of the context".[x] While the SE survey may be suitable for certain uses (i.e., providing formative data to help school leaders set priorities around resources, supports, and practices), it may not be suitable for others (e.g., evaluating the effect of a SE program). As such, school and district leaders should think carefully about whether the CORE Districts' SE survey is suitable for their intended use in their given context. An additional consideration is that the research from Education Analytics shows that using scaled scores (i.e., estimated using Item Response Theory, or IRT, models) can increase the precision of the survey results, by taking into account differences in the amount of information offered by each item and addressing issues related to missing student data. However, we recognize that reporting raw (or non-scaled) responses is often easier for most districts, and therefore provide benchmark data based on non-scaled responses in this report.

Survey Use in Practice

Assessing social-emotional development through student surveys can offer a myriad of benefits to schools and districts. The authors of the Assessment Work Group Practitioner Guide: *Choosing and Using SEL Competency Assessments: What Schools and Districts Need to Know* suggest that measuring student SE competencies should be used, among other things, to include student voice as a component in schools' decision-making process, communicate SEL as a priority, deepen educators' understanding of how SE competencies develop in students over time, improve SE instruction and implementation through a continuous improvement model, and help support equitable outcomes in education.[xi]

Educators in the CORE Districts are using SE data to pave the way for changes in practice unique to their districts and schools. For example, in Los Angeles Unified School District (LAUSD), social-emotional learning facilitators help teachers and administrators review student survey data and design strategic plans to more intentionally integrate SEL throughout the school day. Additionally, schools in San Francisco Unified School District have used SE survey data to create district-wide SEL standards and then used those standards to incorporate teacher reports of student SEL into student report cards.

There are several considerations for practitioners when administering SE surveys and interpreting SE survey data. First and foremost, SE assessments are only one piece of a balanced assessment system for understanding the factors that improve students' social-emotional development. For example, school administrators might also look at school climate data and teacher perceptions of students' SE development in order to get a more complete picture. Further, there are several ways in which the survey data may produce inaccurate, misleading, or even biased results. For example, SE surveys may be susceptible to reference bias, which can occur when students in one context respond to the survey differently from students in a different context based upon the cultural norms that they observe. While this was not found to be an issue in the CORE districts writ large, states and districts using this instrument should investigate this phenomena to ensure results aren't biased.[xii] Research also suggests that student survey results tend to decline from fall to spring and across grade-levels, which suggests caution in using the data to evaluate the effects of single and multi-year policies and programs[xiii]. Finally, given that students can indicate any response of their choosing on the survey, the survey results should not be attached to any type of stakes, such as a high-stakes accountability system, in which the survey results may be corrupted.

Interpreting CORE Districts' Benchmarking Data

This section provides benchmark data from the CORE districts' social-emotional survey administered in 2015-16 to nearly half a million 4th-12th grade students, calculated by Education Analytics. The CORE benchmarking is particularly useful given that the SEL research base is still nascent, and there is still much to learn about how students perceive and report on SE competencies. Classroom teachers, school leaders, and district and state administrators can use benchmark data to better understand students' SE development in order to determine which resources and practices can better support all students. As such, SE survey data can be disaggregated by grade level, race/ethnicity, gender, and special education status and compared with CORE Districts benchmarking data in order to enable educators to examine student strengths, needs, and disparities within a school. Exploring the data in this way can enable schools to hypothesize and initiate data-informed conversations about how to change policies, practices, and school culture to better serve all students, especially underserved students.

Researchers have begun examining trends from the CORE district benchmark data. In one study, based on two years of data, researchers produced a snapshot of student performance across grade levels to illuminate trends among students as they transition through middle and high school.[xiv] The authors found that, with the exception of growth mindset, self-reported SE skills do not increase steadily as students move through school in the same way academic skills do. Rather, SE skills dip in middle school years and partially recover in high school years. Further, there are differences in self-reported scores across student subgroups that appear to be more than simple aberrations. For example, although female students report slightly higher self-management and social awareness compared to male students, their perceptions of their self-efficacy drop sharply relative to male students in middle school.

Understanding that SE data do not follow the same trajectory and should not be interpreted in the same way as students' academic test score data is critical for utilizing these benchmark data. Two considerations are particularly important to emphasize:

- Students' self-reported scores year after year may reflect changes in students' internal standards about what it means to strongly demonstrate a particular competency *in addition* to changes in their underlying competencies. As such, students who score lower than the CORE average may not necessarily be lacking a particular competency, but rather have higher standards for that particular skill (as discussed above).
- Raw or unscaled scores of student results (i.e., a 3 on a 5-point Likert scale) do not take into account naturally occurring trends in reporting across grades or across scales. Therefore, it is not appropriate to compare raw scores across grade-levels or scales (note that such comparisons would be enabled by using psychometric approaches for constructing a "vertical scale," such that scores in one grade could be directly compared to scores in a different grade). Instead, comparing aggregate scores by subgroup and grade level to the CORE benchmark data and taking into account the distribution of the CORE benchmark data by subgroup and grade level can offer insight into whether students within a school, district, or state have social-emotional competencies and mindsets that are above or below average for a given grade. Ongoing research at Education Analytics aims to construct a vertically equated scale that can then enable more precise comparisons of scores across grade levels.

Tables 1-4 contain benchmark data across the four SE constructs based on data from the CORE Districts in 2015-16. Each table displays the average raw scores of student perceptions of their self-reported growth mindset, self-efficacy, self-management, and social awareness by grade level (grades 4-12) and demographic subgroup (gender, race/ethnicity, ELL status, FRL status, and special education status). Standard deviations are also included, which measure the variability or spread around the mean. These data are based on raw scores reported on a 5-point Likert scale.

In order to make inferences about their students' social-emotional competency and mindset development, users can compare their aggregated data across grade levels and subgroups for a given construct to the appropriate CORE benchmark data. When interpreting a school or district's data, TransformEd recommends the following standards: 1 standard deviation above the CORE benchmark average indicates above-average scores, 1 standard deviation below the CORE benchmark average indicates below-average scores, and scores within 1 standard deviation of the mean indicates average scores. For example, in Table 1, among all fourth-grade students, the average growth mindset score was 3.7, and the standard deviation was 1.2. If a given school, district, or state's fourth-grade average score is 2.4, it indicates that their students have a growth mindset that is one standard deviation below the CORE benchmark data, or "below average", and could thus indicate the need to implement practices geared toward improving students' growth mindset.

Table 1. The CORE Districts, School Year 2015-16: Growth Mindset Results by Grade Level and Student Demographics

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
All Students	Mean	3.7	3.7	3.7	3.5	3.4	3.3	3.3	3.3	3.4
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Gender										
Female	Mean	3.7	3.7	3.7	3.4	3.3	3.2	3.2	3.2	3.3
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Male	Mean	3.6	3.7	3.7	3.6	3.5	3.4	3.4	3.5	3.5
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)
Race/Ethnicity										
Asian	Mean	3.7	3.7	3.8	3.7	3.7	3.5	3.4	3.3	3.3
	Std Dev	(1.1)	(1.3)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)
Black	Mean	3.8	3.8	3.8	3.6	3.5	3.4	3.4	3.4	3.5
	Std Dev	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
Hispanic	Mean	3.6	3.6	3.6	3.4	3.3	3.3	3.3	3.3	3.4
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Other	Mean	3.8	3.8	3.8	3.7	3.6	3.4	3.5	3.4	3.4
	Std Dev	(1.2)	(1.1)	(1.1)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)	(1.2)
White	Mean	3.9	4.0	4.0	3.9	3.8	3.6	3.5	3.5	3.6
	Std Dev	(1.1)	(1.1)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
English Language Learner (ELL) Status										
ELL	Mean	3.5	3.4	3.4	3.3	3.2	3.2	3.2	3.2	3.3
	Std Dev	(1.3)	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.2)
Not ELL	Mean	3.8	3.8	3.7	3.6	3.5	3.4	3.3	3.4	3.4
	Std Dev	(1.2)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Free-and-Reduced-Price Lunch (FRL) Status										
FRL	Mean	3.6	3.6	3.6	3.4	3.4	3.3	3.3	3.3	3.4
	Std Dev	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)
Not FRL	Mean	3.9	3.9	3.9	3.8	3.7	3.5	3.4	3.4	3.5
	Std Dev	(1.1)	(1.1)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Special Education (SPED) Status										
SPED	Mean	3.4	3.4	3.4	3.3	3.2	3.1	3.2	3.2	3.3
	Std Dev	(1.4)	(1.3)	(1.3)	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
Not SPED	Mean	3.7	3.7	3.7	3.5	3.5	3.4	3.3	3.4	3.4
	Std Dev	(1.2)	(1.2)	(1.1)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)

Table 2. The CORE Districts, School Year 2015-16: Self-Efficacy Results by Grade Level and Student Demographics

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
All Students	Mean	3.7	3.7	3.7	3.5	3.4	3.3	3.3	3.3	3.4
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Gender	Mean									
Female	Std Dev	3.7	3.7	3.7	3.4	3.3	3.2	3.2	3.2	3.3
		(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Male	Mean	3.6	3.7	3.7	3.6	3.5	3.4	3.4	3.5	3.5
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)
Race/Ethnicity										
Asian	Mean	3.7	3.7	3.8	3.7	3.7	3.5	3.4	3.3	3.3
	Std Dev	(1.1)	(1.3)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Black	Mean	3.8	3.8	3.8	3.6	3.5	3.4	3.4	3.4	3.5
	Std Dev	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
Hispanic	Mean	3.6	3.6	3.6	3.4	3.3	3.3	3.3	3.3	3.4
	Std Dev	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Other	Mean	3.8	3.8	3.8	3.7	3.6	3.4	3.5	3.4	3.4
	Std Dev	(1.2)	(1.1)	(1.1)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)	(1.2)
White	Mean	3.9	4.0	4.0	3.9	3.8	3.6	3.5	3.5	3.6
	Std Dev	(1.1)	(1.1)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
English Language Learner (ELL) Status										
ELL	Mean	3.5	3.4	3.4	3.3	3.2	3.2	3.2	3.2	3.3
	Std Dev	(1.3)	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.2)
Not ELL	Mean	3.8	3.8	3.7	3.6	3.5	3.4	3.3	3.4	3.4
	Std Dev	(1.2)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Free-and-Reduced-Price Lunch (FRL) Status										
FRL	Mean	3.6	3.6	3.6	3.4	3.4	3.3	3.3	3.3	3.4
	Std Dev	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)
Not FRL	Mean	3.9	3.9	3.9	3.8	3.7	3.5	3.4	3.4	3.5
	Std Dev	(1.1)	(1.1)	(1.1)	(1.1)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)
Special Education (SPED) Status										
SPED	Mean	3.4	3.4	3.4	3.3	3.2	3.1	3.2	3.2	3.3
	Std Dev	(1.4)	(1.3)	(1.3)	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
Not SPED	Mean	3.7	3.7	3.7	3.5	3.5	3.4	3.3	3.4	3.4
	Std Dev	(1.2)	(1.2)	(1.1)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)

Table 3. The CORE Districts, School Year 2015-16: Self-Management Results by Grade Level and Student Demographics

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
All Students	Mean	4.0	4.1	4.1	4.0	4.0	4.0	4.1	4.1	4.1
	Std Dev	(1.1)	(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)
Gender	Mean									
Female	Std Dev	4.2	4.2	4.2	4.1	4.1	4.1	4.1	4.2	4.2
		(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)
Male	Mean	3.9	4.0	4.0	4.0	4.0	4.0	4.0	4.1	4.1
	Std Dev	(1.2)	(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)
Race/Ethnicity										
Asian	Mean	4.2	4.2	4.3	4.3	4.2	4.2	4.2	4.2	4.1
	Std Dev	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)	(0.9)	(0.9)	(0.9)	(0.9)
Black	Mean	3.9	3.9	4.0	3.9	3.9	3.9	4.0	4.1	4.1
	Std Dev	(1.2)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)	(1.0)
Hispanic	Mean	4.0	4.0	4.1	4.0	4.0	4.0	4.0	4.1	4.2
	Std Dev	(1.1)	(1.1)	(1.0)	(1.1)	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)
Other	Mean	4.1	4.2	4.2	4.1	4.1	4.1	4.1	4.2	4.1
	Std Dev	(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(0.9)	(1.0)
White	Mean	4.3	4.4	4.4	4.3	4.3	4.2	4.2	4.2	4.2
	Std Dev	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)	(0.9)	(0.9)	(0.9)	(1.0)
English Language Learner (ELL) Status										
ELL	Mean	3.8	3.9	3.9	3.8	3.8	3.8	3.9	4.0	4.0
	Std Dev	(1.2)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.0)	(1.1)
Not ELL	Mean	4.1	4.2	4.2	4.1	4.1	4.0	4.1	4.2	4.2
	Std Dev	(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Free-and-Reduced-Price Lunch (FRPL) Status										
FRL	Mean	4.0	4.0	4.1	4.0	4.0	4.0	4.1	4.1	4.2
	Std Dev	(1.1)	(1.1)	(1.0)	(1.1)	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)
Not FRL	Mean	4.2	4.3	4.3	4.3	4.2	4.1	4.1	4.2	4.2
	Std Dev	(1.0)	(1.0)	(1.0)	(0.9)	(1.0)	(1.0)	(1.0)	(0.9)	(1.0)
Special Education (SPED) Status										
SPED	Mean	3.7	3.7	3.8	3.8	3.8	3.8	3.9	3.9	4.0
	Std Dev	(1.3)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)
Not SPED	Mean	4.1	4.1	4.2	4.1	4.0	4.0	4.1	4.2	4.2
	Std Dev	(1.1)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(0.9)	(0.9)

Table 4. The CORE Districts, School Year 2015-16:, Social Awareness Results by Grade Level and Student Demographics

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
All Students	Mean	3.9	3.9	3.8	3.7	3.6	3.5	3.6	3.6	3.7
	Std Dev	(1.1)	(1.1)	(1.0)	(1.1)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)
Gender	Mean									
Female	Std Dev	4.1	4.0	3.9	3.7	3.6	3.6	3.6	3.7	3.7
		(1.1)	(1.0)	(1.0)	(1.0)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)
Male	Mean	3.8	3.8	3.8	3.6	3.5	3.5	3.5	3.6	3.6
	Std Dev	(1.1)	(1.1)	(1.0)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)
Race/Ethnicity										
Asian	Mean	3.9	3.9	3.8	3.7	3.7	3.7	3.7	3.7	3.7
	Std Dev	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)

		Grade Level								
		4	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Black	Mean	3.8	3.8	3.7	3.6	3.5	3.5	3.5	3.6	3.6
	Std Dev	(1.2)	(1.2)	(1.1)	(1.1)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)
Hispanic	Mean	3.9	3.9	3.8	3.6	3.5	3.5	3.5	3.6	3.6
	Std Dev	(1.1)	(1.1)	(1.0)	(1.1)	(1.1)	(1.1)	(1.1)	(1.0)	(1.1)
Other	Mean	4.0	4.0	3.9	3.7	3.6	3.7	3.7	3.7	3.7
	Std Dev	(1.1)	(1.0)	(1.0)	(1.0)	(1.1)	(1.0)	(1.0)	(1.0)	(1.1)
White	Mean	4.0	4.1	4.0	3.9	3.8	3.7	3.7	3.7	3.8
	Std Dev	(1.0)	(1.0)	(0.9)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)
English Language Learner (ELL) Status										
ELL	Mean	3.8	3.8	3.7	3.6	3.5	3.5	3.5	3.5	3.6
	Std Dev	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)	(1.1)
Not ELL	Mean	4.0	3.9	3.9	3.7	3.6	3.6	3.6	3.6	3.7
	Std Dev	(1.1)	(1.0)	(1.0)	(1.0)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)
Free-and-Reduced-Price Lunch (FRPL) Status										
FRL	Mean	3.9	3.9	3.8	3.6	3.5	3.5	3.5	3.6	3.6
	Std Dev	(1.1)	(1.1)	(1.0)	(1.1)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)
Not FRL	Mean	4.0	4.0	3.9	3.8	3.7	3.6	3.6	3.7	3.7
	Std Dev	(1.0)	(1.0)	(1.0)	(1.0)	(1.1)	(1.0)	(1.0)	(1.0)	(1.0)
Special Education (SPED) Status										
SPED	Mean	3.7	3.7	3.6	3.5	3.5	3.4	3.5	3.5	3.5
	Std Dev	(1.3)	(1.2)	(1.2)	(1.2)	(1.2)	(1.1)	(1.1)	(1.1)	(1.2)
Not SPED	Mean	4.0	3.9	3.9	3.7	3.6	3.6	3.6	3.6	3.7
	Std Dev	(1.1)	(1.0)	(1.0)	(1.0)	(1.1)	(1.1)	(1.1)	(1.0)	(1.0)

Uses of Benchmarking Data in Practice: Examples from the Field

The CORE benchmark data are intended to provide information to help teachers, school leaders, and district/state administrators make sense of their survey data regarding student SE competencies and make data-informed decisions. Such decisions can ensure that educators and administrators are providing and allocating appropriate resources and supports to ensure the proper SE development of all students. See PACE's working paper, which provides additional insight on the practices, interventions, and supports being offered in schools with strong social-emotional development. [xv]

Transforming Education's partnership with NewSchools Venture Fund (NewSchools) and the NewSchools Invent portfolio of schools is a prime example of the use of CORE districts' benchmark data to inform practice.[xvi] The Invent portfolio consists of over 60 innovative schools throughout the country that are dedicated to supporting an expanded definition of student success, which includes a focus on student social-emotional development alongside academic performance and the culture and climate of the school.[xvii] Through this partnership, TransformEd oversees the administration of a broader set of measures, incorporating the CORE scales to schools in the Invent portfolio. TransformEd then provides coaching to school leaders on how to interpret the data and translate the results into actionable steps. This data coaching utilizes CORE benchmarking data to help school leaders explore and make sense of student perceptions of their SE competencies. This is done, in part, through the use of school and cohort dashboards that include CORE districts' grade and subgroup-level benchmarks in growth mindset, self-efficacy, self-management, and social awareness. Inclusion of these benchmarks in the Invent dashboards allow exploration of school- and cohort-specific SE data of schools in the Invent portfolio relative to students in the same grade and demographic subgroups within the CORE districts. School leaders in the Invent cohort are using their school-specific dashboards to identify bright spots and surprises in the patterns and trends relative to other schools in the Invent portfolio and the CORE districts.

Based on our partnership with the NewSchools Invent portfolio, as well as our work with districts throughout the country that are making use of the CORE benchmarking data, we offer three ways that practitioners can use benchmark data to target resources and supports needed most within their schools and districts. First, benchmark data can be used to illuminate strengths in particular schools or grade-levels in order to help identify and scale promising practices. Second, benchmark data can help leaders and administrators identify disparities in SE development in order to inform resource allocation. Third, benchmark data can be used to prioritize student SE development goals and set priorities for the year.

Identify and scale promising practices

Using benchmark data can help educators and administrators illuminate the strengths of subgroups, classrooms, grade-levels, and schools. By comparing scores to the benchmarking data at the subgroup/grade-level and considering the classroom and school environment, educators can develop hypotheses regarding factors that strengthen students' SE development and identify ways to scale promising practices. For instance, the benchmarking data reveal that female students' self-efficacy tends to drop in the middle school years. However, suppose in one

particular middle school, female students report consistently high self-efficacy. This could signal to district leadership to investigate the utilized practices and supports being offered in that middle school which may be useful to scale to other middle schools in the district.

Identify subgroup disparities for resource allocation

It is important for schools to ensure that every student has access to the supports and resources they need to develop and strengthen their SE skills. Comparing school or district trends to those from the CORE benchmarking data can illuminate discrepancies in the supports received by students across subgroups. For example, CORE benchmarking data reveal that social-emotional competency gaps between socio-economically disadvantaged students and their peers tend to narrow in high school. Finding that gaps in the district are widening in high school can indicate the need to devote additional resources and supports to socio-economically disadvantaged students to ensure they have the competencies necessary to enroll and persist in college.

Prioritize specific goals toward student SE development

Interpreting school or district data on SE competencies along with CORE benchmark data can empower educators to identify specific goals around social-emotional development and set school or district-wide priorities. Suppose that one school has an ongoing SEL program with a specific focus on improving self-efficacy and self-management in their middle school students. Based on results from the school's SE data over time, it is evident that there have been improvements in both self-management and self-efficacy. However, when compared to the CORE benchmarking data, there still exists a considerable gap between the school's data and the CORE Districts' benchmarking data in terms of male students' self-management. Leaders might choose to set the goal of improving male students' self-management and make intentional efforts to integrate elements of self-management practices into teacher instruction and professional development.

Conclusion

Student SE survey data can offer important insight into students' social-emotional development by providing information that are not available from other academic and non-academic measures by highlighting gaps in perceptions among subgroups of students within schools. Benchmarking data can further inform the interpretation of survey data by illuminating student strengths and weaknesses with regard to their social-emotional development, which in turn can help administrators determine which supports, practices, and resources should be implemented in order to improve and strengthen student outcomes.

As research on the CORE SEL survey progresses, additional resources for making sense of and interpreting student data from the CORE SEL survey will continue to emerge. Although the benchmarks provided here provide a useful starting point for the field as practitioners and policymakers work to interpret SEL data, there are additional technical issues that warrant consideration before the data provided here can be considered reliable, valid benchmarks. In particular, efforts to construct a vertically equated scale (i.e., a scale that is comparable across grades) and a horizontally equated scale (i.e., a scale that is comparable across multiple years of administration) are necessary to accurately allow for meaningful and accurate comparisons across grades and years. Ongoing research by Education Analytics aim to move in this direction to further contribute to actionable information for the field.

Given the increasing administration of SE surveys in elementary and secondary education across the United States, it is imperative that educators use and interpret SE survey data correctly and that they be aware of the strengths and limitation of the various uses of SE survey data in classrooms, schools, and districts. To that end, state, district, and school leaders should establish systems and structures that support the use and interpretation of SE data and integrate relevant benchmark data to inform classroom practice and changes in curriculum.[xviii] School leaders and administrators should ensure that educators have access to clear guidance on the appropriate use of SE data. Lastly, school and district leaders should ensure access to high-quality professional learning opportunities around using SE survey data to improve educators' understanding of students' SE competencies and their development over time.

[i] National Commission on Social, Emotional, and Academic Development. (2019). From A Nation at Risk to a Nation at Hope. Retrieved from http://nationathope.org/wp-content/uploads/2018_aspen_final-report_full_webversion.pdf

[ii] Gehlbach, H and Hough, H. (2018). Measuring Social Emotional Learning through Student Surveys in the CORE Districts: A Pragmatic Approach to Validity and Reliability. *Policy Analysis for California Education.* Retrieved from https://www.edpolicyinca.org/publications/sel-validity

[iii] Measurement of SE skills are based on a student self-report survey, which features a set of items designed to measure the four SE competencies: self-management (9 items), social awareness (8 items), growth mindset (4 items), and self-efficacy (4 items). See https://coredistricts.org/wp-content/uploads/2017/07/SEL_6.9.17.docx for the full list of items and a description of the scoring approach.

[iv] Krachman, S.B., Arnold, R. & LaRocca, B. (2016). Expanding the Definition of Student Success: A Case Study of the CORE Districts. *Transforming Education*. Retrieved from https://www.transformingeducation.org/wp-content/uploads/2017/04/TransformingEducationCaseStudyFINAL1.pdf

[v] Hough, H, Kalogrides, D., & Loeb S. (2017). Using Surveys of Students' Social-Emotional Skills and School Climate for Accountability and Continuous Improvement. *Policy Analysis of California Education.* Retrieved from https://www.edpolicyinca.org/publications/using-sel-and-cc

[vi] Gehlbach, H and Hough, H. (2018). Measuring Social Emotional Learning through Student Surveys in the CORE Districts: A Pragmatic Approach to Validity and Reliability. *Policy Analysis for California Education.* Retrieved from https://www.edpolicyinca.org/publications/sel-validity

[vii] West, M.R., Buckley, K., Krachman, S.B., and Bookman, N. (2018). Development and Implementation of Student Social-Emotional Surveys in the CORE Districts. *Journal of Applied Developmental Psychology, 55.* Retrieved from https://www.sciencedirect.com/science/article/pii/S0193397316301290

[viii] Meyer, R.H., Wang, C., & Rice, A.B. (2018). Measuring Students' Social-Emotional Learning Among California's CORE Districts: An IRT Modeling Approach. *Policy Analysis for California Education.* Retrieved from https://edpolicyinca.org/sites/default/files/Measuring_SEL_May-2018.pdf

[ix] Hough, H, Kalogrides, D., & Loeb S. (2017). Using Surveys of Students' Social-Emotional Skills and School Climate for Accountability and Continuous Improvement. *Policy Analysis of California Education.* Retrieved from https://www.edpolicyinca.org/publications/using-sel-and-cc

[x] Gehlbach, H and Hough, H. (2018). Measuring Social Emotional Learning through Student Surveys in the CORE Districts: A Pragmatic Approach to Validity and Reliability. *Policy Analysis for California Education,* page 23. Retrieved from https://www.edpolicyinca.org/publications/sel-validity

[xi] Taylor, J.J., Buckley, K., Hamilton, L.S., Stecher, B.M., Read, L., & Schweig, J. (2018). Choosing and Using SEL Competency Assessments: What Schools and Districts Need to Know. *Measuring SEL: Using Data to Inspire Practice.* Retrieved from http://measuringsel.casel.org/pdf/Choosing-and-Using-SEL-Competency-Assessments_What-Schools-and-Districts-Need-to-Know.pdf

[xii] West, M. (2017). Should Non-Cognitive Skills Be Included in School Accountability Systems? Preliminary Evidence from California's CORE Districts. *Brookings.* Retrieved from https://www.brookings.edu/research/should-non-cognitive-skills-be-included-in-school-accountability-systems-preliminary-evidence-from-californias-core-districts/

[xiii] Buckley, K. & Krachman, S.B. Patterns in Student Self-Report and Teacher Report Measures of Social-Emotional Mindsets, Skills, and Habits. *Transforming Education.* Retrieved from https://www.transformingeducation.org/wp-content/uploads/2017/04/TE-BCRCWorkingPaperFINAL.pdf

[xiv] West, M.R., Pier, L., Fricke, H., Hough, H., Loeb, S., Meyer, R.H., & Rice, A.B. (2018). Trends in Student Social Emotional Learning: Evidence from the CORE Districts. *Policy Analysis for California Education.* Retrieved from https://edpolicyinca.org/publications/sel-trends

[xv] Marsh, J.A., McKibben, S., Hough, H., Hall, H., Allbright, T.N., Matewos, A.N., & Siqueira, C. (2018). Enacting Social-Emotional Learning: Practices and Supports Employed in CORE Districts and Schools. *Policy Analysis for California Education.* Retrieved from https://www.edpolicyinca.org/publications/sel-practices

[xvi] See https://www.transformingeducation.org/bridging-research-practice-to-expand-the-definition-of-student-success/ and https://www.newschools.org/wp-content/uploads/2018/08/NSVF_080318_EDSS_Insight_Brief_003-rev.pdf for more information on this partnership.

[xvii] See NewSchools Venture Fund Framework: https://www.newschools.org/wp-content/uploads/2019/02/EDSS_slide_standard-w-logo.pdf

[xviii] Hough, H., Byun, E., & Mulfinger, L. (2018). Using Data for Improvement: Learning from the CORE Data Collaborative. *Getting Down to Facts II*. Retrieved from https://www.edpolicyinca.org/publications/using-data-improvement-learning-core-data-collaborative