

School Differences in Social–Emotional Learning Gains: Findings From the First Large-Scale Panel Survey of Students

Susanna Loeb

Brown University

Michael S. Christian

Education Analytics

Heather Hough

Policy Analysis for California Education

Robert H. Meyer

Andrew B. Rice

Education Analytics

Martin R. West

Harvard University

Measures of school-level growth in student outcomes are common tools for assessing the impacts of schools. The vast majority of these measures use standardized tests as the outcome of interest, even though emerging evidence demonstrates the importance of social–emotional learning (SEL). In this article, we present results from using the first large-scale panel surveys of students on SEL to produce school-level value-added measures by grade for growth mindset, self-efficacy, self-management, and social awareness. We found substantive differences across schools in SEL growth, with magnitudes of differences similar to those for growth in academic achievement. In contrast, we found that the goodness of fit of the value-added model was considerably lower when the outcome variables were measures of SEL constructs rather than of academic achievement. In addition, the across-school variance in the average level of the SEL measures was proportionally much smaller than that for academic measures. These findings recommend caution in interpreting measures as the causal impacts of schools on SEL, though they also do not rule out important school effects.

Keywords: *school growth; value-added models; social–emotional learning*

State departments of education, as well as many school districts, use growth measures as tools to assess the impacts of schools and teachers. By one account, growth is used to measure school performance in 42 of the 50 U.S. states and in the District of Columbia (Thomsen, 2013). Much of the literature on growth measures, which include value-added or academic progress measures, has analyzed teacher-level growth (e.g., Chetty, Friedman, & Rockoff, 2014; Rivkin, Hanushek, & Kain, 2005). However, a substantial number of studies have also addressed school-level value-added measures. Topics covered by these studies include the conceptualization and estimation of school effects (Ehlert, Koedel, Parsons, & Podgursky, 2016; Meyer, 1997; Raudenbush & Willms, 1995; Rardon & Raudenbush, 2009; Tekwe et al., 2004), the implications of school growth measures in accountability systems (Kane & Staiger, 2002; Ladd & Walsh, 2002), the persistence of school value-added over time (Briggs & Weeks, 2011), the usage and adaptation of school value-added measures as tools to evaluate the impacts of principals (Chiang, Lipscomb, & Gill, 2016; Grissom, Kalogrides, & Loeb, 2015), and tests of the validity of school effects using data from school-choice lotteries (Angrist, Hull, Pathak, & Walters, 2016, 2017; Deming, 2014).

The vast majority of studies of growth in education have focused on outcomes in academic subjects, such as in mathematics and English language arts (ELA), based on student performance on standardized tests. This focus on academic subjects persists despite a substantial body of emerging research that has found that social–emotional skills (sometimes called noncognitive skills) contribute to school success and adult outcomes (Heckman & Rubenstein, 2001; Kautz, Heckman, Diris, ter Weel, & Borghans, 2014). Cohen, Garcia, Apfel, and Master (2006), for example, found that a brief in-class writing assignment affirming sense of personal adequacy significantly improved the grades of African American students and reduced the racial achievement gap. Similarly, Blackwell, Trzesniewski, and Dweck (2007) found that an intervention teaching on incremental theory of intelligence (“growth mind-set”) to seventh graders increased reported classroom motivation and grades. These and other studies have demonstrated that school performance depends on more than the knowledge and skills typically measured by standardized tests.

While many factors contribute to students’ social–emotional skills, research has increasingly provided evidence that experiences in schools can affect social–emotional learning (SEL) both directly (Allensworth & Easton, 2007; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011) and through the implementation of policies and practices that improve a school’s culture and climate and promote positive relationships (Battistich, Schaps, & Wilson, 2004; Berkowitz, Moore, Astor, & Benbenishty, 2016; Blum, Libbey, Bishop, & Bishop, 2004; Hamre & Pianta, 2006; P. A. Jennings & Greenberg, 2009; McCormick, Cappella, O’Connor, & McClowry, 2015). A meta-analysis by Durlak, Weissberg, Dymnicki, Taylor, and Schellinger (2011) found that school programs and

interventions, such as the ones studied by Cohen et al. (2006) and Blackwell et al. (2007), can improve social–emotional skills. Kautz, Heckman, Diris, ter Weel, and Borghans (2014) noted that the short follow-up of most studies of elementary school programs makes it difficult to draw strong conclusions about their long-term impacts (particularly in comparison to early childhood programs). They also noted that evidence from the studies that have followed participants into adulthood is promising.

Recent studies also showed that teachers can affect student social–emotional development (e.g., Gershenson, 2016; J. L. Jennings & DiPrete, 2010; Ladd & Sorensen, 2017; Liu & Loeb, 2018). For example, Ruzek, Domina, Conly, Duncan, and Karabenick (2015) demonstrated teacher effects on academic motivation, while Blazar and Kraft (2017) found effects on self-reported self-efficacy and happiness, and Jackson (2018) found teacher effects on a composite measure of student grade point average (GPA), on-time grade completion, suspensions, and full-day attendance. Teachers who increase test performance are not necessarily the same as those who help students improve their social–emotional skills. In fact, the correlations between teachers' effects on test scores and teachers' effects on nontest scores are weak (Jackson, 2018; Liu & Loeb, 2018). More generally, a large portion of teacher effects on student long-term outcomes, such as college attendance, is not explained by teacher effects on student achievement, suggesting that good teachers not only increase students' test scores but also impact other outcomes (Chamberlain, 2013).

While some research has assessed teacher effects, none to date, of which we are aware, have assessed the extent to which schools at large vary in their students' SEL trajectories. Yet school-level differences, beyond differences across teachers, could impact student development of these skills. Studies have provided evidence that school leaders affect student learning through mechanisms such as building a sense of community that could also affect students' social–emotional development (see Leithwood, Louis, Anderson, & Wahlstrom, 2004; Hallinger, 2005; Waters, Marzano, & McNulty, 2003, for meta-analyses). Moreover, the prevalence of bullying and other culture and climate characteristics of schools can affect students and their social–emotional health and development (Olweus, 1994), and school-based interventions can affect these cultural characteristics (Tofi & Farrington, 2011).

Given the importance of SEL and the evidence of an effect of schooling on SEL, the possibility of measuring differences in growth across schools in SEL outcomes offers a prospect for a fuller understanding of school differences. In so doing, it may provide a more complete framework for identifying areas of improvement and need within schools.

The value-added model (VAM) is a common approach for measuring growth in academic subjects at the school level. The goal of this approach is to adjust differences in student performance across schools for nonschool factors and for prior schooling sufficiently well, so that the estimated school effects measure the

contributions of schools to student academic outcomes and not the differences in which students each school serves. Toward this goal, the VAM framework predicts students' current academic outcomes using data on lagged academic outcomes and other student characteristics. The school value added, a school-level estimated effect, is then the school mean difference between the actual and predicted outcomes for its students (Angrist et al., 2016, 2017; Chiang et al., 2016; Deming, 2014; Ehlert et al., 2016; Meyer & Dokumaci, 2015).

We used this same VAM framework to produce school-level value-added measures by grade for four dimensions of SEL (growth mind-set, self-efficacy, self-management, and social awareness) using the first large-scale survey panel of students with SEL outcomes. The value-added measures covered the growth of more than 150,000 students in Grades 4 through 8 across schools in the California Office to Reform Education (CORE) districts, a group of large urban districts in California that administered annual surveys with items related to SEL beginning in spring 2015 as part of a multiple-measures accountability system under a No Child Left Behind (NCLB) Flexibility Request.¹ We measured student-level SEL outcomes using student responses to SEL-related items in the 2014–2015 and 2015–2016 administrations of the surveys, which we scaled using item response theory (IRT) methods. Given that these are the first two administrations of such a survey at a large scale, this was the first opportunity to measure differences in student growth in SEL outcomes across a variety of schools. For comparison, we also measured school value-added measures in academic assessment scores in mathematics and ELA for the same students.

We found substantive differences across schools in growth in student SEL outcomes. Over the four SEL constructs and five grades, the estimated standard deviation of impacts across schools relative to the standard deviation of SEL outcome measures across students was between .09 and .24. This magnitude was similar to the estimated standard deviation of school effects on academic achievement in math and ELA, which was between .11 and .18 across grades.

In contrast, we found that the covariates in the VAMs, which included lagged SEL measures, lagged math and ELA assessments, and student demographics, explained far less of the variance across students in SEL outcome measures than they explained across students in academic assessment scores. This lack of explanatory power was true not only for the overall variance across both students and schools but also for the within-school, across-student variance. In addition, while we measured differences across schools in growth in SEL outcomes that are comparable in magnitude to those in growth in academic measures, the across-school variance in the average *level* of the SEL measures was proportionally much smaller than the across-school variance in the average level of the academic measures.

The analyses featured in this article, with 2 years of data, represent the first opportunity to analyze change over time in SEL and school differences in that change. They are a first pass at estimating the impacts of attending individual schools on SEL outcomes at a large scale. As additional years of SEL data from

the CORE survey become available, our understanding of SEL school growth measures will improve. For example, incorporating the third survey year when it is available will make it possible to measure the stability of school SEL growth measures from one growth year (2014–2015 to 2015–2016) to another (2015–2016 to 2016–2017). In addition, research employing the results of the CORE survey will be used to improve the CORE survey itself, as part of a process of continuous improvement as in Davidson et al. (2018).

Data

The data for this study came from participating CORE districts in California. The CORE districts together serve more than 1 million students and nearly 20% of the students in California. The central data set included responses to surveys by students in five participating CORE districts (Fresno, Long Beach, Los Angeles, San Francisco, and Santa Ana) in the spring of the 2014–2015 and 2015–2016 school years. The surveys included items addressing four dimensions of SEL: growth mind-set, self-efficacy, self-management, and social awareness. The surveys included between 4 and 9 items for each of the four constructs (see Appendix A in the online version of the article for a list of all survey items). Each item included up to five possible responses indicating a student's report of either the extent of agreement with a statement or the extent of participation in an activity or experience.

West, Buckley, Krachman, and Bookman (2018) described the four SEL constructs as follows:

Growth mindset is the belief that one's abilities can grow with effort. Students with a growth mindset see effort as necessary for success, embrace challenges, learn from criticism, and persist in the face of setbacks (Dweck, 2006). Self-efficacy is the belief in one's own ability to succeed in achieving an outcome or reaching a goal. Self-efficacy reflects confidence in the ability to exert control over one's motivation, behavior, and environment (Bandura, 1997). Self-management is the ability to regulate one's emotions, thoughts, and behaviors effectively in different situations. This includes managing stress, delaying gratification, motivating oneself, and setting and working toward personal and academic goals (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2005). Finally, social awareness is the ability to take the perspective of and empathize with others from diverse backgrounds and cultures, to understand social and ethical norms for behavior, and to recognize family, school, and community resources (CASEL, 2005).

Transforming Education (2016) and West, Buckley, et al. (2018) described the process of developing and introducing the SEL surveys. The survey development team curated and adapted questions from a wide range of researcher-developed, free-to-administer measures of SEL, in consultation with experts in the SEL field. Eighteen schools piloted the survey in 2013–2014, and all district schools field-tested the survey in 2014–2015. Beginning in 2015–2016, the

districts included the survey as part of the system to assess school performance established by the CORE districts' NCLB waiver and included the average levels of the SEL constructs by school (not the growth in these measures) in their dashboard of multiple school-level measures. Since passage of the Every Student Succeeds Act, the successor to NCLB, the districts have used the SEL survey data for informational purposes, not for direct school accountability.

Complementing the data from the student SEL survey were data from the Smarter Balanced Assessment Consortium's (SBAC) assessments in math and ELA, which students in Grades 3 through 8 completed across California. The SBAC assessment is a computer-adaptive assessment aligned to the Common Core standards. The state administered these in the spring 2014–2015 and 2015–2016, allowing us to compare growth in SEL to growth in math and ELA achievement. Because the SBAC is administered in the spring of Grades 3 through 8, we could only measure growth, which requires both a current outcome measure and a prior outcome measure in math and ELA, for students in Grades 4 through 8. As a result, we could compare SEL growth measures to more traditional academic growth measures only in Grades 4 through 8 and so focused on SEL growth in these grades for this study.

The samples we used in producing the SEL growth measures comprised students in CORE districts, who responded to the survey in both 2014–2015 and 2015–2016 (see Appendix B in the online version of the article for an extended description of how we constructed the samples we used). Students must have responded to at least half of the survey questions associated with a given SEL construct for their responses to have been considered valid. To have been included in the growth measure for a given SEL construct, students must have had valid survey responses in 2015–2016 for that particular construct, as well as valid responses in 2014–2015 for all four constructs. We required valid responses to all four constructs in 2014–2015 because all four are control variables in the growth model. In some districts, two forms of the survey were administered; the sample only included students who responded to the more commonly administered form. In addition, students must have had SBAC scores in math and ELA in 2014–2015, have had demographic data available to serve as additional control variables in the growth model, and have been matched to a school in the five participating CORE districts. Similarly, we estimated the SBAC growth measures for a given subject using a sample of students in the five participating CORE districts with SBAC scores in that subject in 2015–2016, SBAC scores in both subjects in 2014–2015, valid responses in all four SEL constructs in 2014–2015, and available demographic data.

Table 1 describes the students in the sample. Panel A of Table 1 characterizes all students in the sample with growth measures in self-efficacy. The demographic makeup of the samples for the other three SEL constructs and for math and ELA was similar; 70% to 75% of the sample was Hispanic, about 7% was Black, and 4% to 9% was Asian. Approximately 80% were eligible for subsidized lunch. Thirty-seven percent of fourth graders were English language learners, but this number dropped to 15% by eighth grade.

TABLE 1.
Descriptive Statistics

(A) Demographics (Proportions of Students by Characteristic, Self-Efficacy Outcome Sample)					
	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	.50	.50	.50	.50	.50
Asian	.04	.05	.08	.09	.08
Hispanic	.75	.73	.71	.70	.70
Black	.07	.07	.07	.06	.07
ELL	.37	.29	.25	.17	.15
Disability	.10	.11	.11	.10	.10
Economically disadvantaged	.81	.80	.81	.79	.79
Foster	.01	.01	.01	.01	.00
Homeless	.01	.02	.02	.02	.03

(B) Number of Students and Schools (Samples for All Four Outcomes)					
	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Number of students					
ELA	35,485	42,399	41,466	38,700	39,203
Math	35,464	42,379	41,426	38,672	39,122
Growth mind-set	31,169	37,749	33,914	31,622	31,987
Self-efficacy	31,244	37,812	33,926	31,642	32,015
Self-management	31,290	37,855	34,002	31,725	32,076
Social awareness	30,903	37,557	33,800	31,476	31,908
Number of schools					
ELA	540	623	351	203	204
Math	540	623	351	203	204
Growth mind-set	549	635	356	209	202
Self-efficacy	549	635	356	209	202
Self-management	549	635	357	209	202
Social awareness	548	635	355	209	202

Note. ELL = English language learners; ELA = English language arts.

Panel B of Table 1 presents the number of students and schools in the sample for each grade and for each outcome variable. There were fewer students for the SEL measures than for math and ELA. This smaller sample was in part a result of nonresponse or incomplete responses to the survey. Differences in the number of students across grades were in part the result of differences in participation over the five sample districts in the SEL survey and SBAC assessment. As shown in Appendix B in the online version of the article, we included substantially overlapping but not identical samples for each SEL construct and academic subject

within each grade to use as many observations as possible to measure academic and SEL growth at the school level. Table A1 in additional tables of demographic characteristics can be found in Appendix C in the online version of the article.

Measuring SEL growth using these data required us to transform the responses to the SEL items on the student survey into a metric. We created scale scores for each of the four SEL constructs for students who responded to at least half of the survey items associated with that construct.² We used a generalized partial credit model (GPCM) to produce a scale score for each of the four constructs from the responses to these items. We estimated a separate GPCM for each construct. Based on Muraki's (1992) extension of the partial credit model (Masters, 1982), GPCM can incorporate measures for which responses are on a multipoint scale in contrast to dichotomous items.³

Meyer, Wang, and Rice (2018); West, Buckley, et al. (2018); West, Pier, et al. (2018); and Gehlbach and Hough (2018) described the properties of the SEL scale score measures. Meyer et al. (2018), which focused on the reliability of the SEL measures, found that the model fit of the GPCM model was statistically significantly lower than that of the nominal response model (NRM); however, the differences were small enough that they recommended using the GPCM when scoring the SEL survey for its relative simplicity, while suggesting the NRM as a useful tool for continued research into improving the survey.⁴ When considering the consistency of responses for the measures within each SEL domain, they found that the internal scale reliabilities, measured using Cronbach's α , of the self-efficacy, self-management, and social awareness scales ranged from .76 to .89. They also found in an exploratory factor analysis that survey items associated with the same construct loaded onto the same factors and did not substantially load onto other factors. They did express some concern with the growth mind-set items, which, after review of item category response functions measured using the NRM, did not appear to function well among younger students. Moreover, the internal scale reliability of the Growth Mindset Scale was lower than for the other measures, below .70 in grades below Grade 7. The issues with growth mind-set may have stemmed from the survey items associated with it that were phrased negatively (e.g., "my intelligence is something I can't change very much," as opposed to the self-efficacy item "I can earn an A in my classes"), which younger students may have misunderstood.

West, Buckley, et al. (2018); West, Pier, et al. (2018); and Gehlbach and Hough (2018) focused on the validity of the SEL survey measures. West, Buckley, et al. (2018) found that, at the school level, measures of all four constructs from the CORE SEL survey were positively correlated with GPA, ELA test scores, and math test scores at the elementary, middle, and high school levels. They also found that all four SEL constructs were negatively correlated across middle schools with the absence rate and that three of the four constructs (self-management, growth mind-set, and social awareness) were negatively correlated with the percentage of students who received suspensions. They found that the

within-school, across-student correlations of the SEL measures with GPA and ELA and math test scores were typically smaller than the overall correlations, which is suggestive evidence against the possibility that differences in SEL measures across schools are distorted by reference bias. In addition, they found a substantive positive correlation between student reports and teacher reports of students' self-management and social awareness.

West, Pier, et al. (2018) compared trends in SEL Scale scores from the CORE survey across grades and found that they often conformed to expectations from research on SEL skills. They found that SEL skills, as measured by the survey, did not necessarily increase as children became older. In particular, they found that survey measures of self-management and self-efficacy declined in middle school, in line with previous research on the topic. In contrast, they found that survey measures of social awareness also declined in middle school, which differs from previous research. Other findings of West, Pier, et al. (2018) include that girls had higher survey measures of self-management and social awareness and that economically disadvantaged students had lower survey measures of all four SEL constructs, which are all findings consistent with existing research on SEL skills. Gehlbach and Hough (2018) surveyed evidence on validity of the survey-based SEL measures, including that from the abovementioned papers, and placed it within a detailed and wide-ranging framework of validity.

Table 2 presents the internal scale reliabilities of the SEL scale scores, measured from CORE-wide data using Cronbach's α . For comparison, Table 2 also presents the reliabilities of the computer-adaptive SBAC assessment in mathematics and ELA, computed from CORE-wide data using IRT conditional standard errors of measurement (SEMs). We present the reliabilities using Cronbach's α for the SEL measures and IRT conditional SEMs for the SBAC assessments because these were the measures of error employed when using errors-in-variables regression to estimate the VAM that produced the school growth measures. We used Cronbach's α for the SEL measures because reliabilities based on IRT conditional SEMs were very low for some of the SEL measures, particularly in the lower grades and particularly in growth mind-set, which had the potential to overadjust for measurement error when estimating the errors-in-variables regression. We used IRT conditional SEMs for the SBAC assessments because the SBAC is a computer-adaptive assessment for which Cronbach's α does not apply.

The reliabilities of the SEL measures were lower than the reliabilities of the SBAC measures, regardless of whether they were measured using Cronbach's α or IRT conditional SEMs. This lower reliability may have resulted in part from the small number of items used to produce the SEL measures relative to the achievement measures, although the SEL measures based on more items did not always have higher reliabilities than the SEL measures based on fewer items. The reliabilities of the SEL measures rose with grade, while those of the academic subjects declined slightly.

TABLE 2.
Internal Scale Reliability of SEL and SBAC Scale Scores, 2014–2015 and 2015–2016

	Growth Mind-Set	Self- Efficacy	Self- Management	Social Awareness	ELA	Math
2014–2015						
Grade 3	0.61	0.78	0.80	0.77	0.91	0.93
Grade 4	0.62	0.83	0.83	0.78	0.91	0.93
Grade 5	0.66	0.85	0.84	0.79	0.92	0.91
Grade 6	0.67	0.86	0.86	0.80	0.90	0.91
Grade 7	0.70	0.88	0.88	0.82	0.91	0.89
Grade 8	0.73	0.88	0.87	0.81	0.91	0.90
2015–2016						
Grade 3	0.63	0.78	0.80	0.76	0.92	0.94
Grade 4	0.64	0.83	0.82	0.77	0.92	0.94
Grade 5	0.68	0.86	0.84	0.79	0.93	0.92
Grade 6	0.70	0.87	0.85	0.80	0.91	0.93
Grade 7	0.71	0.88	0.87	0.82	0.92	0.91
Grade 8	0.74	0.88	0.87	0.82	0.92	0.91
No. of items	4	4	9	8	37–42	30–34

Note. Reliability measures were based on Cronbach’s α in the four SEL constructs (growth mind-set, self-efficacy, self-management, and social awareness) and on IRT conditional SEMs in the two academic subjects (English language arts and mathematics). Number of items in SBAC ELA and mathematics assessments were drawn from assessment blueprints. ELA = English language arts; SEL = social–emotional learning; IRT = item response theory; SEM = standard error of measurement; SBAC = Smarter Balanced Assessment Consortium.

The distributions of the SEL scale scores exhibited some evidence of ceiling effects, with a substantial proportion of students having chosen the most affirmative response to every item within a construct. Relatedly, the SEL raw scores displayed some degree of rightward skew, which was substantially mitigated in the transformation to scale scores. Figure 1 presents histograms of the SEL scale scores (IRT θ scores) in Grades 5 and 8 in a CORE-wide sample as examples; histograms for other grades (provided in Appendix C in the online version of the article) were similar. While the ceiling effects exhibited by the scale scores present challenges for looking at changes among individual students, they do not necessarily inhibit the school-level measures that are the focus of this article (Koedel & Betts, 2010).

*Between- and Within-School Variance and Across-Year
Covariance of SEL Measures*

Before proceeding to assessing the measures of each school’s value-added to SEL, we describe the variance across schools in the levels of the SEL measures

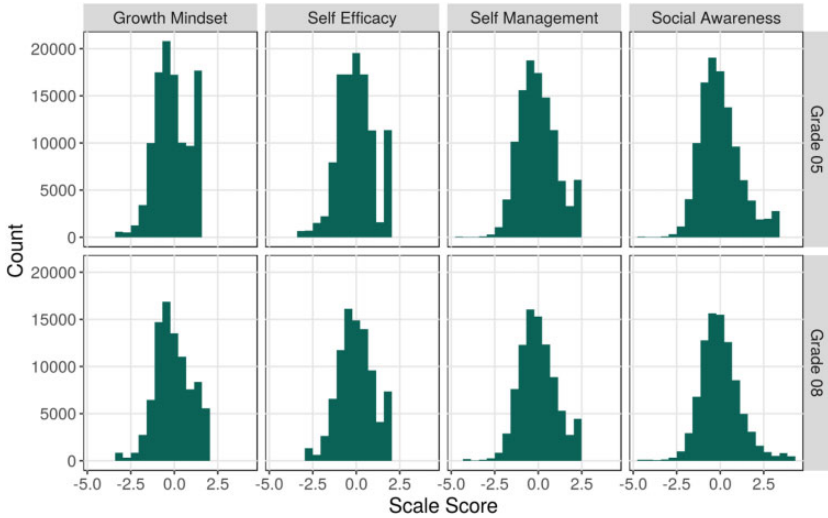


FIGURE 1. Histogram of distribution of Social-Emotional Learning Scale (item response theory θ scores), Grades 5 and 8.

themselves. To understand the across-school and within-school components of both the variance of the scale scores in a given year and the covariance of the scale scores from one year to the next, we estimated the following seemingly unrelated regressions (SUR) model:

$$Y_{cijt} = \mu_{cjt} + \eta_{cijt}, \tag{1}$$

$$Y_{cijt-1} = \mu_{cjt-1} + \eta_{cijt-1}, \tag{2}$$

where j is the school attended by student i in year t , Y_{cijt} is measured scale score in construct c of student i in year τ , μ_{cjt} is the component of the variance of construct c in year τ that is across year- t schools, and η_{cijt} is the component of the variance of construct c in year τ that is within year- t schools. School attended in year t is used to break down variance in both year t and year $t - 1$ to parallel the construction of a school growth model, in which the average growth in scale scores between year $t - 1$ and year t is used to estimate the impact of schools attended in year t .

We estimated Equations 1 and 2 using SUR and from those results obtained estimates of the variances of μ_{cjt} , μ_{cjt-1} , η_{cijt} , and η_{cijt-1} and of the covariances between μ_{cjt} and μ_{cjt-1} and between η_{cijt} and η_{cijt-1} . Note that the variances and covariances involving the student-level η_{cijt} terms include not only variance across students within schools in a construct but also variance from randomness

School Differences in Social–Emotional Learning Gains

TABLE 3.

Across-School and Within-School-Across-Student Components of Variance and Year-to-Year Correlation in Scale Scores in Academic Subjects and SEL Constructs

	Variance of Scale Scores, 2015–2016			Correlation of Scale Scores, 2014–2015 to 2015–2016		
	Total	Across School	Within School	Total	Across School	Within School
Grade 5						
English language arts	1.00	.21	.79	.83	.95	.80
Mathematics	1.00	.22	.78	.85	.93	.83
Growth mind-set	1.00	.10	.90	.33	.65	.30
Self-efficacy	1.00	.05	.95	.43	.56	.42
Self-management	1.00	.07	.93	.53	.84	.51
Social awareness	1.00	.04	.96	.43	.60	.42
Grade 8						
English language arts	1.00	.17	.83	.83	.94	.81
Mathematics	1.00	.19	.81	.84	.95	.82
Growth mind-set	1.00	.03	.97	.44	.90	.43
Self-efficacy	1.00	.03	.97	.53	.81	.52
Self-management	1.00	.05	.95	.55	.89	.53
Social awareness	1.00	.04	.96	.49	.88	.48

with which the assessment or survey measured each student’s academic or SEL outcomes.

Table 3 presents the across-school and within-school variance of the academic subject and SEL construct scale scores in Grades 5 and 8 (results for additional grades can be found in Appendix C in the online version of the article). It also presents the across-school and within-school correlation between current and lagged scale scores in the same grades. Two findings are noteworthy. First, the proportion of the variance in the SEL scale scores that is across school was small in comparison to the same proportion of the variance in the SBAC scale scores. For example, in Grade 5, only 4% of the variance in the social awareness scale score was across school, compared to 22% of the variance in the mathematics scale score. This smaller across-school variation could have been due to greater measurement error or to a smaller school effect. Second, the correlation from year to year in the SEL scale scores was substantially lower than that in the SBAC scale scores for both across-school and within-school components. This pattern suggests that student-level SEL outcomes, as measured by the survey, had lower persistence over time than the academic measures.

Method

Creating Growth Measures

We modeled the impacts of schools in academic subjects and SEL constructs using the following value-added regression model:

$$y_{cijt} = \xi_c + y_{Cijt-1}\lambda_c + X_{ijt}\beta_c + \alpha_{cjt} + \varepsilon_{cijt}, \quad (3)$$

where school j is the school attended by student i in year t , y_{cijt} is outcome in construct c for student i in school j in year t , y_{Cijt-1} is a 1×6 vector of outcomes in all six subjects and constructs (math, ELA, and the four SEL constructs) for student i in year $t-1$, X_{ijt} is a vector of characteristics of student i in year t , α_{cjt} is the impact of school j on growth in construct c in year t , ε_{cijt} is a student error term, and λ_c and β_c are conformable coefficient vectors. This specification has been referred to as a covariate adjustment model (McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004) and as a dynamic ordinary least squares model (Guarino, Reckase, & Wooldridge, 2015).

The growth model presented in Equation 3 was specified with the purpose of isolating the effect of schooling in year t on student i 's academic or SEL outcomes from the effects of schooling in years before t and from the effects of nonschool factors experienced over student i 's entire lifetime up to and including year t . The lagged outcome variables y_{Cijt-1} and student characteristics X_{ijt} were included as covariates in the model to control for the effects of previous schooling and for nonschool factors as much as possible given the data available. More specifically, we included the lagged outcome variables y_{Cijt-1} to adjust for relevant school and nonschool factors experienced up to the end of year $t-1$, the effects of which are likely to have been substantially reflected in student outcomes at the end of year $t-1$. We included the student demographic variables X_{ijt} as adjustments for nonschool factors experienced in year t that may have been correlated with school assignment in that year. Remaining variance in relevant nonschool factors experienced in year t after controlling for lagged outcomes and student demographics will only bias growth measures to the extent that they were correlated with school assignment in year t .

The model in Equation 3 includes an effect at the school level, α_{cjt} , but not at the teacher level. This school effect can be interpreted as combining the average effect of the teachers associated with the school with any effects associated with the school outside of the effects of its teachers. The concept of a school effect that incorporates the effects of the school's teachers is common in practical applications, which typically do not attempt to control for teacher effects when producing school growth measures; see, for example, technical reports for school growth measures for the states of New York (Education Analytics, 2017) and Ohio (SAS, 2018). This approach to modeling school effects is also widely used in research studies of school-level value-added measures (Angrist et al., 2016,

2017; Chiang et al., 2016; Deming, 2014; Ehlert et al., 2016; Meyer & Dokumaci, 2015). Creating an estimated single-year school effect measure that does not include the teachers' effects is not possible because no student attends school for the year without a teacher.

We estimated Equation 3 using the errors-in-variables approach described in Fuller (1987), which uses an estimate of the variance of measurement error in the right-hand side variables to correct the sums of squares and cross products matrix, such that it reflects the variances and covariances of the variables in the model had they not been measured with error. In this application, the variance of measurement error is measured using Cronbach's α for lagged SEL constructs and IRT conditional SEMs for lagged SBAC scores. Given that the right-hand side variables are the same regardless of which outcome is used as the left-hand side variable, it makes no difference whether Equation 3 is estimated separately for each equation or jointly as a system of SUR. The student characteristics (X_{ijt}) included on the right-hand side of the regression were gender, ethnicity, ELL status, economically disadvantaged status, disability status, foster status, and homeless status.

We centered the school fixed effect estimates from this regression to have a weighted mean of zero, with the weight equal to the number of students associated with the school in the regression sample. As a result, the school fixed effects were measured relative to the average school effect across the schools in the sample. We used these centered school effect estimates as the measures of school growth for each of the six outcomes. Both the current and lagged scale scores were rescaled to have a mean of zero and a standard deviation of one within each regression sample, so that the school growth measures were measured in units of standard deviation across students in the outcome scale score measure.

Properties of Growth Measures

We examined four properties of the school growth measures. First, we computed goodness-of-fit measures to assess the extent to which the model's covariates, which included lagged SEL and academic outcomes and student demographics, predicted SEL and academic outcomes. We used traditional R^2 and within-school R^2 to measure model fit. Second, we estimated the variance in school growth, adjusted for sampling error in the estimated school effects, to describe the magnitude of differences in SEL and academic growth across schools. This approach to measuring the variance of school effects adjusts for sampling error but not for other possible forms of measurement error or systematic bias. Third, we measured correlations of the school growth measures across constructs. This analysis examined whether schools in which students gained more than expected in one dimension also gained more than expected in any of the other dimensions. Finally, we

looked at similar correlations for each construct across grades within a school. For each of these four analyses, we looked across SEL measures and also compared the SEL measures to academic achievement measures. As a specification check, we repeated the analyses with growth measures that included school aggregate measures of student characteristics as well as the student-level achievement and SEL measures.

Results

Model Coefficients and Goodness of Fit

Coefficients and goodness-of-fit measures for regression models of academic and SEL outcomes appear in Table 4 (Appendix C in the online version of the article presents tables of the coefficients of all covariates). Since both current and lagged outcome variables were standardized by grade, the coefficients on the lagged outcome measures can be interpreted as the increase in standard deviations of the outcome variable associated with a one standard deviation increase in the lagged outcome variable. For example, in the model in which the outcome variable was self-management in fifth grade, the coefficient on lagged self-management was .50. This implies that a one standard deviation increase in self-management in fourth grade was associated with a half standard deviation increase in self-management in fifth grade. The coefficient of .09 on lagged ELA achievement in the same model implies that a one-standard-deviation increase in ELA achievement in fourth grade was associated with a .09 standard deviation increase in self-management in fifth grade.

In all SEL models, the greatest coefficients were on same-construct lag. All coefficients except one were between .35 and .56. The exception was fourth-grade growth mind-set, for which the coefficient was .23. The largest coefficients on same-construct lag were for social awareness (.43–.56), followed by self-management (.42–.50). Coefficients on same-construct lag were generally lower for growth mind-set (.23–.50) and self-efficacy (.36–.46). All of these coefficients were smaller than the coefficients on the same-subject lag in models of math and ELA achievement, which ranged from .61 (seventh-grade ELA) to .94 (eighth-grade math).

The goodness-of-fit measures were smaller in the models of the SEL measures than in the models of academic achievement. The overall R^2 of the SEL models ranged from .27 (fourth-grade growth mind-set) to .44 (fifth-grade self-management), considerably lower than that of the academic subject models, which ranged from .82 (sixth-grade ELA) to .92 (eighth-grade math). Among the SEL models, the overall R^2 was higher in self-management (between .39 and .44) than in growth mind-set (.27–.41), self-efficacy (.28–.41), or social awareness (.28–.38).

School Differences in Social–Emotional Learning Gains

TABLE 4.
Coefficients and Regression Statistics

(A) Coefficients From Models of SEL Constructs, Grade 4

	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self- Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.20 (0.02)	0.05 (0.02)	0.13 (0.01)	0.09 (0.02)
Math	0.07 (0.01)	0.16 (0.01)	0.08 (0.01)	−0.04 (0.01)
Growth	0.23 (0.01)	0.02 (0.01)	0.00 (0.01)	0.00 (0.01)
mind-set				
Self-efficacy	0.05 (0.01)	0.36 (0.01)	0.00 (0.01)	−0.01 (0.01)
Self-	0.05 (0.01)	0.04 (0.01)	0.42 (0.01)	0.05 (0.01)
management				
Social	0.01 (0.01)	0.01 (0.01)	0.05 (0.01)	0.43 (0.01)
awareness				
<i>N</i>	31,169	31,244	31,290	30,903
<i>R</i> ²	.27	.28	.39	.28
Within-school <i>R</i> ²	.14	.20	.29	.20
School fixed effects	549	549	549	548

(B) Coefficients From Models of SEL Constructs, Grade 5

	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self- Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.17 (0.01)	0.05 (0.01)	0.09 (0.01)	0.07 (0.01)
Math	0.07 (0.01)	0.12 (0.01)	0.06 (0.01)	−0.04 (0.01)
Growth	0.35 (0.01)	0.11 (0.01)	0.05 (0.01)	0.07 (0.01)
mind-set				
Self-efficacy	0.09 (0.01)	0.39 (0.01)	−0.02 (0.01)	−0.04 (0.01)
Self-	0.02 (0.01)	0.05 (0.01)	0.50 (0.01)	0.07 (0.01)
management				

(continued)

TABLE 4. (continued)

(B) Coefficients From Models of SEL Constructs, Grade 5

	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self-Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Social awareness	0.02 (0.01)	0.04 (0.01)	0.06 (0.01)	0.50 (0.01)
<i>N</i>	37,749	37,812	37,855	37,557
<i>R</i> ²	.35	.34	.44	.34
Within-school <i>R</i> ²	.21	.25	.35	.24
School fixed effects	635	635	635	

(C) Coefficients From Models of SEL Constructs, Grade 6

	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self-Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.17 (0.01)	-0.02 (0.01)	0.07 (0.01)	0.02 (0.02)
Math	0.05 (0.01)	0.20 (0.01)	0.07 (0.01)	0.03 (0.01)
Growth mind-set	0.41 (0.01)	0.06 (0.01)	0.05 (0.01)	0.04 (0.01)
Self-efficacy	0.08 (0.01)	0.41 (0.01)	-0.03 (0.01)	-0.04 (0.01)
Self-management	0.01 (0.01)	0.04 (0.01)	0.49 (0.01)	0.03 (0.01)
Social awareness	0.02 (0.01)	0.04 (0.01)	0.07 (0.01)	0.54 (0.01)
<i>N</i>	33,914	33,926	34,002	33,800
<i>R</i> ²	.38	.34	.43	.35
Within-school <i>R</i> ²	.27	.28	.34	.26
School fixed effects	356	356	357	355

(continued)

School Differences in Social–Emotional Learning Gains

TABLE 4. (continued)

(D) Coefficients From Models of SEL Constructs, Grade 7

	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self- Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.11 (0.02)	−0.02 (0.02)	0.04 (0.01)	−0.01 (0.02)
Math	0.10 (0.02)	0.16 (0.02)	0.09 (0.01)	0.05 (0.02)
Growth mind-set	0.48 (0.01)	0.09 (0.01)	0.04 (0.01)	0.05 (0.01)
Self-efficacy	0.07 (0.01)	0.45 (0.01)	−0.01 (0.01)	−0.05 (0.01)
Self- management	0.03 (0.01)	0.05 (0.01)	0.50 (0.01)	0.06 (0.01)
Social awareness	−0.01 (0.01)	0.04 (0.01)	0.09 (0.01)	0.56 (0.01)
<i>N</i>	31,622	31,642	31,725	31,476
<i>R</i> ²	.41	.38	.43	.38
Within-school <i>R</i> ²	.29	.32	.34	.28
School fixed effects	209	209	209	209

(E) Coefficients From Models of SEL Constructs, Grade 8

	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self- Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.14 (0.02)	−0.01 (0.02)	0.08 (0.02)	0.05 (0.02)
Math	0.03 (0.02)	0.19 (0.02)	0.06 (0.02)	0.01 (0.02)
Growth mind-set	0.50 (0.01)	0.11 (0.01)	0.05 (0.01)	0.07 (0.01)
Self-efficacy	0.06 (0.01)	0.46 (0.01)	0.01 (0.01)	−0.03 (0.01)
Self- management	0.03 (0.01)	0.04 (0.01)	0.48 (0.01)	0.05 (0.01)
Social awareness	0.04 (0.01)	0.03 (0.01)	0.09 (0.01)	0.55 (0.01)
<i>N</i>	31,987	32,015	32,076	31,908

(continued)

TABLE 4. (continued)

(E) Coefficients From Models of SEL Constructs, Grade 8				
	Current Outcome (Left-Hand Side Variable)			
	Growth Mind-Set	Self-Efficacy	Self-Management	Social Awareness
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
R^2	.41	.41	.41	.38
Within-school R^2	.30	.35	.34	.29
School fixed effects	202	202	202	202

(F) Coefficients From Models of ELA and Math, Grades 4 and 5				
	Current Outcome (Left-Hand Side Variable)			
	ELA, Grade 4	Math, Grade 4	ELA, Grade 5	Math, Grade 5
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.68 (0.01)	0.07 (0.01)	0.70 (0.01)	0.11 (0.01)
Math	0.16 (0.01)	0.80 (0.01)	0.16 (0.01)	0.79 (0.01)
Growth mind-set	0.01 (0.01)	0.01 (0.01)	0.03 (0.01)	0.01 (0.01)
Self-efficacy	0.01 (0.01)	0.02 (0.01)	0.01 (0.01)	0.02 (0.00)
Self-management	0.06 (0.01)	0.04 (0.01)	0.05 (0.01)	0.03 (0.01)
Social awareness	-0.04 (0.01)	-0.04 (0.01)	-0.03 (0.01)	-0.03 (0.01)
N	35,485	35,464	42,399	42,379
R^2	.82	.85	.85	.86
Within-school R^2	.71	.75	.74	.76
School fixed effects	540	540	623	623

(continued)

School Differences in Social–Emotional Learning Gains

TABLE 4. (continued)

(G) Coefficients From Models of ELA and Math, Grades 6 and 7

	Current Outcome (Left-Hand Side Variable)			
	ELA, Grade 6	Math, Grade 6	ELA, Grade 7	Math, Grade 7
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged outcomes				
ELA	0.71 (0.01)	0.16 (0.01)	0.61 (0.01)	0.03 (0.01)
Math	0.12 (0.01)	0.70 (0.01)	0.26 (0.01)	0.89 (0.01)
Growth	0.02 (0.01)	0.00 (0.01)	0.04 (0.01)	0.00 (0.00)
mind-set				
Self-efficacy	0.02 (0.00)	0.03 (0.00)	0.00 (0.00)	0.02 (0.00)
Self-	0.06 (0.01)	0.06 (0.01)	0.04 (0.01)	0.03 (0.01)
management				
Social	−0.04 (0.01)	−0.05 (0.01)	−0.02 (0.01)	−0.02 (0.01)
awareness				
<i>N</i>	41,466	41,426	38,700	38,672
<i>R</i> ²	.82	.85	.83	.92
Within-school <i>R</i> ²	.73	.76	.73	.81
School fixed effects	351	351	203	203

(H) Coefficients From Models of ELA and Math, Grade 8

	Current Outcome (Left-Hand Side Variable)	
	ELA, Grade 8	Math, Grade 8
	Coefficient (SE)	Coefficient (SE)
Lagged outcomes		
ELA	0.66 (0.01)	0.00 (0.01)
Math	0.24 (0.01)	0.94 (0.01)
Growth	0.04 (0.00)	0.00 (0.01)
mind-set		
Self-efficacy	−0.01 (0.00)	0.01 (0.00)
Self-	0.02 (0.00)	0.03 (0.01)
management		
Social	0.01 (0.00)	0.00 (0.01)
awareness		

(continued)

TABLE 4. (continued)

	Current Outcome (Left-Hand Side Variable)	
	ELA, Grade 8	Math, Grade 8
	Coefficient (SE)	Coefficient (SE)
<i>N</i>	39,203	39,122
<i>R</i> ²	.85	.92
Within-school <i>R</i> ²	.76	.78
School fixed effects	204	204

Note. All regressions included lagged outcomes in ELA, math, growth mind-set, self-efficacy, self-management, and social awareness; indicators for gender, economic disadvantage, English language learner (beginning, intermediate, advanced, and level not measured), foster child, homeless, race/ethnicity (Asian, Black, Hispanic), and disability (moderate, severe); and school fixed effects. ELA = English language arts; *SE* = standard error.

The *R*² measures described above include not only the explanatory power of student characteristics, such as lagged achievement, lagged SEL-related skills, and demographics, but also the explanatory power of the school fixed effects. In order to distinguish the explanatory power of the individual controls from that of the school fixed effects, we estimated a within-school *R*², which measures the extent to which the outcome variable is predicted specifically by the variables that are included in the model toward the goals of controlling for nonschool factors and identifying school effects (Gawade & Meyer, 2016). As was the case in overall goodness of fit, the within-school measure was substantially lower in the SEL models than in the models of academic achievement. Across the four SEL outcomes and five grades, within-school *R*² ranged between .14 (fourth-grade growth mind-set) and .35 (eighth-grade self-efficacy). In contrast, the same combination of explanatory variables explained about three quarters of the within-school variation in math and ELA achievement.

The lower within-school *R*² in the estimated models of SEL outcomes means that the variables included in the model to control for nonschool factors did not have the same predictive power for SEL outcomes as they did for academic outcomes. This finding could have been the result of greater measurement error in the outcome variable, which would artificially inflate the variation of the measure of the SEL constructs. However, it could also be suggestive of the possibility that these control variables may not have sufficiently controlled for nonschool factors in the SEL models, which, if these factors are correlated with school assignment, would lead to omitted-variables bias in the SEL growth measures.

Variance and Reliability of School Growth Estimates

We estimated the overall magnitude of the impacts of schools with the noise-corrected variance of estimated school effects in the SEL growth models as follows:

$$\text{Est. Variance}[\alpha_{cjt}] = \text{Sample Variance}[\hat{\alpha}_{cjt}] - \text{Sample Mean}[\hat{\sigma}_{cjt}^2], \quad (4)$$

where α_{cjt} , as in the notation of Equation 3, are the effects for school j in construct c at time t , $\hat{\alpha}_{cjt}$ are the estimated school effects produced by the value-added regression and centered to have a mean of zero, and $\hat{\sigma}_{cjt}^2$ are the squares of the standard error estimates of those estimated and centered effects. This approach estimates the variance of the component of the school effects α_{cjt} in a way that does not include variance due to the estimation error of $\hat{\alpha}_{cjt}$.

Using this estimate of the variance of α_{cjt} across schools, we also estimated the average reliability of the growth measure estimates as follows:

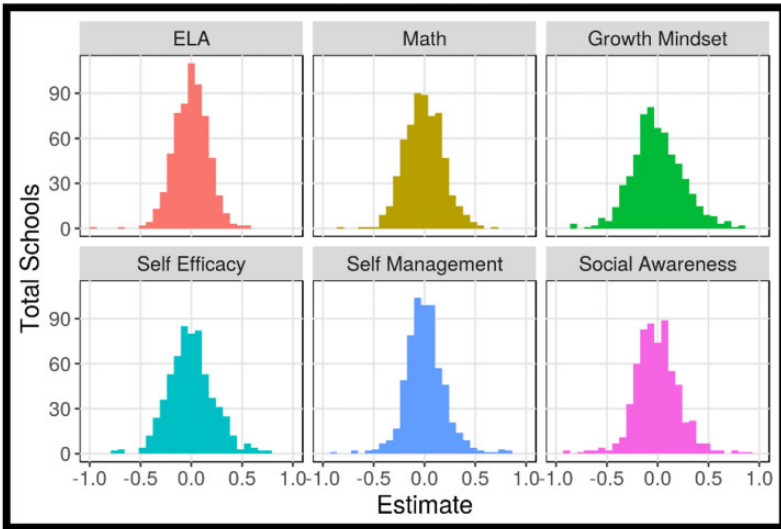
$$\text{Avg. Reliability} [\hat{\alpha}_{cjt}] = \text{Est. Variance}[\alpha_{cjt}] / \text{Sample Variance}[\hat{\alpha}_{cjt}]. \quad (5)$$

The average reliability is the proportion of the variance of the estimated school effects $\hat{\alpha}_{cjt}$ that is the result of variance of the underlying school effect α_{cjt} and not the result of variance in the error with which $\hat{\alpha}_{cjt}$ estimates α_{cjt} . This reliability measure includes as estimation error the effect of randomness in student SEL growth on school SEL growth measures. It does not, however, include in estimation error any potentially distortionary school-level factors (such as, e.g., the weather on the day the survey is taken) that might affect measures of SEL growth at the school level. A high-reliability school growth measure captures differences in effects across schools with more precision and less randomness.

Figure 2 presents histograms of the estimated school effects for each of the six outcomes for Grades 5 and 8; results in other grades (presented in Appendix C in the online version of the article) were similar. The histograms illustrate that the range of growth measures across schools was very similar for the four SEL constructs in comparison to the two academic subjects. Moreover, the distributions were approximately normal.

Table 5 presents estimates of the standard deviation of school effects in models of growth in both SEL constructs and academic subjects. The estimated standard deviation of school effects on mathematics and ELA was between .11 and .18 times the standard deviation of achievement across students, depending on subject and grade. The magnitudes were smaller, although not especially so, than the standard deviation of about .20 estimated in Angrist, Hull, Pathak, and Walters (2017) in their study of school effects in sixth-grade math in Boston.⁵ In contrast, the magnitudes were larger than Kane and Staiger's (2001) estimates of the standard deviation of fifth-grade school gains in math (.15) and reading (.08) in North Carolina.

Panel A: Grade 5



Panel B: Grade 8

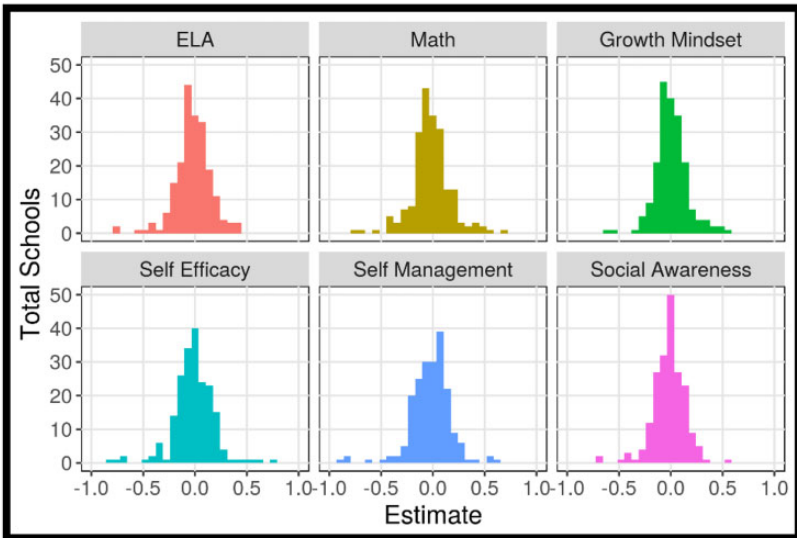


FIGURE 2. Histograms of school effects.

TABLE 5.
Standard Deviation (Noise-Corrected) and Reliability of School Growth Effect Estimates

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Standard deviation					
ELA	.15	.14	.18	.13	.13
Math	.17	.16	.16	.11	.13
Growth mind-set	.24	.23	.15	.11	.09
Self-efficacy	.18	.19	.12	.10	.11
Self-management	.16	.13	.13	.11	.11
Social awareness	.16	.16	.13	.11	.10
Reliability					
ELA	.84	.84	.93	.92	.93
Math	.89	.89	.93	.92	.92
Growth mind-set	.80	.80	.73	.69	.61
Self-efficacy	.69	.75	.67	.70	.73
Self-management	.67	.61	.71	.75	.75
Social awareness	.64	.65	.67	.71	.66

Note. ELA = English language arts.

The results in Table 5 show that the variance of school effects in models of growth in SEL constructs was similar to the variance of school effects in models of growth in academic subjects. For example, in fifth grade, the estimated standard deviation across schools of effects on growth mind-set was equal to .23 times the standard deviation of growth mind-set outcomes across students. Across all five grades, the estimated standard deviation of school effects was in the range of .09 and .24 across the four SEL constructs. The standard deviation tends to be lower in middle school grades than in elementary school grades across all four SEL constructs and both academic subjects.

Correlations of School Growth Measures Across Constructs

Table 6 presents the correlations between school growth measures across the two academic and four SEL outcomes. These correlations, as well as the correlations in Tables 7 through 9, are presented as properties of the estimated growth measures $\hat{\alpha}_{cjt}$ and include both the covariance among the underlying school impact parameters α_{cjt} and the variance or covariance in the error with which $\hat{\alpha}_{cjt}$ estimates α_{cjt} . The correlations within the academic measures and within the SEL measures were larger than those between the academic and SEL measures. These differences are not unexpected. Prior work on teachers has shown that teachers who excel in improving students' academic achievement did not necessarily improve other outcomes for students (Jackson, 2018; Liu & Loeb, 2018), and the same logic may apply to school effects.

TABLE 6.
Correlation of Academic and SEL School Growth Measures

	ELA	Math	Growth Mind-Set	Self- Efficacy	Self- Management	Social Awareness
Grade 4						
ELA	1.00	.69	.17	.08	.15	.15
Math	.69	1.00	.16	.14	.19	.14
Growth mind-set	.17	.16	1.00	.40	.26	.39
Self-efficacy	.08	.14	.40	1.00	.55	.61
Self-management	.15	.19	.26	.55	1.00	.67
Social awareness	.15	.14	.39	.61	.67	1.00
Grade 5						
ELA	1.00	.63	.23	.13	.16	.18
Math	.63	1.00	.20	.09	.14	.13
Growth mind-set	.23	.20	1.00	.49	.30	.38
Self-efficacy	.13	.09	.49	1.00	.55	.54
Self-management	.16	.14	.30	.55	1.00	.61
Social awareness	.18	.13	.38	.54	.61	1.00
Grade 6						
ELA	1.00	.74	.39	.15	.26	.26
Math	.74	1.00	.43	.15	.27	.27
Growth mind-set	.39	.43	1.00	.25	.20	.31
Self-efficacy	.15	.15	.25	1.00	.51	.48
Self-management	.26	.27	.20	.51	1.00	.70
Social awareness	.26	.27	.31	.48	.70	1.00
Grade 7						
ELA	1.00	.43	.30	.04	.22	.25
Math	.43	1.00	.18	.08	.12	.23
Growth mind-set	.30	.18	1.00	.37	.34	.35
Self-efficacy	.04	.08	.37	1.00	.47	.43
Self-management	.22	.12	.34	.47	1.00	.68
Social awareness	.25	.23	.35	.43	.68	1.00
Grade 8						
ELA	1.00	.43	.22	.18	.23	.27
Math	.43	1.00	.25	.13	.18	.31
Growth mind-set	.22	.25	1.00	.29	.27	.18
Self-efficacy	.18	.13	.29	1.00	.47	.34
Self-management	.23	.18	.27	.47	1.00	.64
Social awareness	.27	.31	.18	.34	.64	1.00

Note. ELA = English language arts.

TABLE 7.
Correlations of School Effects Across Grades Within Schools

	Elementary Grades			Middle Grades		
	4 and 5	5 and 6	4 and 6	6 and 7	7 and 8	6 and 8
ELA	.20	.16	.24	.25	.41	.17
Math	.10	.12	.06	.00	.27	.20
Growth mind-set	.17	.05	.05	.25	.30	.11
Self-efficacy	.18	.05	.12	.22	.28	.24
Self-management	.09	.04	.11	.33	.47	.32
Social awareness	.10	−.02	−.16	.27	.37	.37

Note. ELA = English language arts.

Correlations of School Growth Measures Within Schools Across Grades

Table 7 presents the correlations of school growth measures across grades within schools. The table includes only correlations within elementary grades and within middle grades, with sixth grade included as both an elementary and a middle grade. Fewer than 50 schools in the CORE sample included both elementary and middle grades.

The correlations among school SEL growth measures across grades were modest, but they were similar to the correlations across grades among the academic growth measures; the average of the correlations presented in Table 7 was .18 among the SEL constructs and also .18 among the academic subjects. These modest correlations suggest that there was substantial variation in the impacts of schools on both academic and SEL outcomes by grade. This pattern would be consistent with the presence of variance in effects across individual teachers and classrooms within schools. The within-school, across-grade correlation was generally greater in the middle grades than in the elementary grades for both the SEL and academic outcomes; this result potentially stems from teachers having taught in multiple grades within middle schools and students having experienced multiple teachers each year.

Sensitivity of School Growth Measures to Including School-Level Covariates

The VAM employed throughout this article included only student-level variables among its covariates. An alternative VAM would control not only for student-level variables but also for school-level variables. One version of this model, which controls for school-level averages of all variables included as student-level covariates, is described in Equations 6 and 7:

$$y_{cijt} = \xi_c + \lambda_c y_{Cijt-1} + X_{ijt} \beta_c + \alpha_{cjt} + \varepsilon_{cijt}, \tag{6}$$

$$\alpha_{cjt} = \zeta_c + \theta_c \bar{y}_{Cjt-1} + \bar{X}_{jt} \phi_c + u_{cjt}, \tag{7}$$

TABLE 8.
Correlations Between Value-Added Measures That Include and Do Not Include Controls for School-Level Averages of Student-Level Covariates

Grade	4	5	6	7	8
ELA	.95	.98	.96	.88	.91
Mathematics	.95	.96	.97	.90	.95
Growth mind-set	.98	.98	.95	.89	.90
Self-efficacy	.95	.93	.94	.92	.86
Self-management	.97	.96	.91	.85	.85
Social awareness	.98	.96	.94	.86	.85

Note. ELA = English language arts.

where Equation 6 is the same as Equation 3, \bar{y}_{Cjt-1} is a vector of average prior academic and SEL outcomes at school j , \bar{X}_{jt} is a vector of average student demographics at school j , and u_{cjt} is the value-added effect of school j . The model expressed in Equations 6 and 7 has the advantage of controlling for any school-level peer effects that are correlated with average prior academic and SEL outcomes or demographics. It has the disadvantage of partialing out from value added any aspect of school quality that is correlated with these school-level averages. The model can be estimated using a two-step approach. First, the regression in Equation 6 is estimated using an errors-in-variables regression, as described earlier in the discussion of Equation 3, to produce the school fixed effects estimates $\hat{\alpha}_{cjt}$. Second, the school fixed effects estimates $\hat{\alpha}_{cjt}$ are regressed on the school-level averages \bar{y}_{Cjt-1} and \bar{X}_{jt} by ordinary least squares as in Equation 7. The residual from the second regression are estimates of the value-added effects u_{cjt} . Alternatively, the model can be estimated in a single step by regressing the school-level averages of the left-hand side variable in Equation (6) on \bar{y}_{Cjt-1} and \bar{X}_{jt} using ordinary least squares. The residuals from this regression will estimate the value-added effects u_{cjt} .

Table 8 presents correlations between estimated value-added measures between the model used in the bulk of this article, which only controlled for student-level covariates, and a model that also controlled for school-level averages of the student-level covariates. For the most part, the correlations were high, ranging from .85 to .99. Even when correlations were high, however, the value-added measures of individual schools may have differed substantially between the two approaches.

The properties of the SEL school growth measures produced by the model that controlled for school averages were, for the most part, similar to those of the model that controlled for student-level variables only. The fit of the model was not substantially improved by the inclusion of the school-level averages.⁶ The estimated variances of both the academic and SEL school effects were slightly

TABLE 9.
Correlations Between Middle Grades Within Schools Between Models That Control and Do Not Control for School Averages

	Model Without School Averages			Model With School Averages		
	6 and 7	7 and 8	6 and 8	6 and 7	7 and 8	6 and 8
ELA	.25	.41	.17	.26	.43	.21
Math	.00	.27	.20	.04	.22	.15
Growth mind-set	.25	.30	.11	.22	.20	.09
Self-efficacy	.22	.28	.24	.25	.18	.14
Self-management	.33	.47	.32	.21	.33	.22
Social awareness	.27	.37	.37	.14	.22	.26

lower in the model that controlled for school-level averages, which is an expected result given that the component of school effects that was correlated with school averages was partialled out. However, the variances were not lower in a way that was sufficiently disproportionate to change the result that the variance of the school effects in the SEL constructs was of a similar magnitude to the variance of the school effects in academic subjects. After having controlled for school averages, the estimated standard deviations of school effects were in the range of .08 to .24 among the SEL constructs and in the range of .11 to .17 among the academic subjects.

One substantive difference between the properties of the growth measures produced by the two models was the correlations among effects within schools across the middle grades. We present a comparison of these correlations in Table 9. The correlations were somewhat lower in the model that included school averages, especially for self-management and social awareness, suggesting that a part of the correlation among estimated effects within schools across the middle grades was driven by the component of those effects, which was correlated with observable student characteristics.

Conclusion

Using data from a large-scale survey panel of more than 150,000 students in five California districts that includes items measuring SEL outcomes, we produced and evaluated measures of the impacts of individual schools on SEL outcomes by grade. To our knowledge, this was the first attempt to produce school growth measures of social–emotional outcomes at a large scale.

The student surveys included items relevant to four SEL outcomes: growth mind-set, self-efficacy, self-management, and social awareness. We used measures of these four SEL outcomes based on responses to the student survey as outcome variables in value-added growth models for Grades 4 through 8. We

estimated the VAMs using linear errors-in-variables regressions of current SEL outcome on lagged SEL outcomes, lagged math and ELA achievement, student demographics, and school fixed effects, with adjustments for measurement error in all lagged SEL and achievement measures. The specification of this value-added growth model was similar to that often used to measure the impacts of schools in academic subjects such as math and ELA, which we also estimated for schools in the districts administering the survey.

We found variance across schools in measured impacts on SEL outcomes, which was similar to the estimated variance across schools in impacts on academic outcomes. Across the four SEL outcomes and five grades covered by this study, we estimated a standard deviation of school effects in the range of .09 and .24 times the standard deviation of the level of SEL outcome measures across students. The analogous standard deviation estimates in models of math and ELA were in the range of .11 to .18.

In contrast, the fit of the VAMs of SEL outcomes was relatively low compared to VAMs of mathematics and ELA. While the covariates in the VAMs explained about three quarters of the variation in math and ELA achievement across students within schools, they did not typically explain more than a third of the variation in the SEL measures. In addition, the SEL measures had more variation within schools than between schools relative to the academic measures.

The extent to which the SEL growth measures and the models underlying them were similar to or different from their analogues for academic achievement is not necessarily an indication of their validity as measures of school impacts on SEL outcomes. That said, given the newness of these measures and the indications of potential measurement issues, we recommend interpreting with caution the value-added measures as measures of the causal effects of schools on students' SEL.

One potential avenue for investigating the extent to which these SEL growth measures reflect causal impacts is to observe how growth measures produced in ordinary circumstances predict these measures in situations in which school assignment can be plausibly believed to be random. This approach has been used to investigate school academic growth measures using school lotteries (Angrist et al., 2016, 2017; Deming, 2014). Blazar (2018) used this approach to investigate SEL survey growth measures at the teacher level using variation from a study in which students within schools were randomly assigned to teachers.

School-level SEL growth measures can be useful even when we do not know the extent to which these growth measures capture causal effects. The results above show substantive differences in measured SEL growth from school to school, which has the potential to help identify schools in which SEL outcomes are falling behind their expected trajectory and thereby help identify where resources such as training or staffing related to SEL may be best allocated. Including school SEL growth measures in a school accountability framework may also be useful, even if they are not causal estimates because they may align

school incentives toward SEL and broaden the emphasis in the school accountability framework beyond strictly academic outcomes. Blazar (2018) discussed these ideas.

Inclusion of SEL growth measures in accountability systems has potential drawbacks. If the growth measures do not capture the causal effects of schools on students' SEL, as they may not, then educators may be frustrated to be held accountable for factors that are out of their control. Moreover, including the SEL growth measures in an accountability system may provide incentives for schools to “coach” students to answer the surveys in a way that produces high growth measures, which would bias not only the growth measures but also the individual student outcome measures. These potential issues recommend substantial care in using survey-based SEL growth measures in a high-stakes policy context.

The validity of the SEL growth measures for schools depends not only on whether the estimates are causal but also on the validity of the SEL survey items that underlie the measures. Surveys are not the only way to measure SEL outcomes. Outcomes such as suspension rates and chronic absenteeism, which are included in CORE's dashboard of measures relevant to school performance, can also be used as proxies for SEL. These alternative measures have the benefit of not relying on student self-reports, though they also have the disadvantage of not clearly distinguishing the particular SEL skills of interest.

The school SEL growth measures described in this study were based on 2 years of CORE student survey data about student SEL outcomes—the minimum sufficient for measuring growth and, to our knowledge, the first panel data set of this size of its kind. As more years of data become available, it will become possible to explore additional issues, including the stability of SEL growth measures for individual schools from year to year, as well as to continue to explore the potential to distinguish the effects of schools on students' SEL in further depth. In addition, continued research on SEL outcomes will inform the evolving design of the CORE survey and of the SEL measures, which will affect the school SEL growth measures in turn. Given the newness of the data, it is most appropriate to understand these results as a first pass at understanding the potential for measuring the impacts of individual schools on SEL outcomes.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported in this presentation was supported by the Walton Family Foundation (Grant # 2017-1553). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Walton Family Foundation.

Notes

1. Fresno, Long Beach, Los Angeles, Oakland, San Francisco, and Santa Ana Unified School Districts were part of the NCLB waiver, and Garden Grove and Sacramento City Unified School Districts are also part of the CORE network.
2. Changing the definition of a valid response to include only students who responded to all survey items for a given construct yielded school growth measures with correlations of between .85 and .96 with the set of SEL school growth measures used in this study. The differences between these school growth measures were in part a result of the sample becoming smaller when the more restrictive all-items criterion was applied (the sample reduced to between 52% and 72% of its previous size). As the size of the sample becomes smaller, the extent to which the variance of measured school growth reflects randomness in growth across the remaining students becomes greater.
3. Using a partial credit model (PCM) in place of the generalized partial credit model (GPCM) to produce SEL scale scores yielded very similar school growth measures, with correlations of .998 or greater depending on grade and construct. Using raw scores also yielded very similar growth measures, with correlations between .985 and .999. An important advantage of using scales based on GPCM (or the PCM), as opposed to raw scores, is that the model provides consistent estimates of scale scores for students with some missing responses, given the assumption that nonresponse to an item is random given the true individual scale—the assumption of local independence.
4. An advantage of the GPCM over the nominal response model that is specific to measuring growth using the social-emotional learning (SEL) survey is that it imposes ordering on the responses to the survey items. This ordering ensures that changes over time in a student's responses to the survey questions are translated into changes in scale scores in a way that reflects how the survey items were designed to measure improvements and/or declines in self-assessed SEL skills.
5. This standard deviation was for a time-invariant effect measured across multiple years rather than a time-variant effect measured in a single year. That stated, results in the appendix of Angrist et al. (2017) implied that the standard deviation of a time-variant effect was of a magnitude of similar size.
6. Both overall R^2 and within-school R^2 are algebraically the same between the two models, and so neither measure sheds much light on the relative fit of the two models. An alternative measure of fit would compare the variance of the explained component minus the school effect $\hat{\xi}_c + \hat{\lambda}_c Y_{Cijt-1} + \hat{\beta}_c X_{ijt} + \hat{\theta}_c \bar{Y}_{Cjt-1} + \hat{\phi}_c \bar{X}_{jt}$ to the variance of the error component plus the school effect $\hat{u}_{cjt} + \hat{\varepsilon}_{cijt}$. The ratio of these two was not substantially different from the analogous ratio in the model that did not control for school averages among the SEL construct or academic subject models.

References

- Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Chicago, IL: Consortium on Chicago School Research.
- Angrist, J., Hull, P., Pathak, P., & Walters, C. (2016). Interpreting tests of school VAM validity. *American Economic Review: Papers and Proceedings*, *106*, 388–392.
- Angrist, J., Hull, P., Pathak, P., & Walters, C. (2017). Leveraging lotteries for school value-added: Testing and estimation. *Quarterly Journal of Economics*, *132*, 871–919.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.
- Battistich, V., Schaps, E., & Wilson, N. (2004). Effects of an elementary school intervention on students' "connectedness" to school and social adjustment during middle school. *Journal of Primary Prevention*, *24*, 243–262.
- 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*, *87*, 425–469.
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*, 246–263.
- Blazar, D. (2018). Validating teacher effects on students' attitudes and behaviors: Evidence from random assignment of teachers to students. *Education Finance and Policy*, *13*, 281–309.
- Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis*, *39*, 146–170.
- Blum, R. W., Libbey, H. P., Bishop, J. H., & Bishop, M. (2004). School connectedness—Strengthening health and education outcomes for teenagers. *Journal of School Health*, *74*, 231–235.
- Briggs, D., & Weeks, J. (2011). The persistence of school-level value-added. *Journal of Educational and Behavioral Statistics*, *36*, 616–637.
- Chamberlain, G. E. (2013). Predictive effects of teachers and schools on test scores, college attendance, and earnings. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 17176–17182.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*, 2633–2679.
- Chiang, H., Lipscomb, S., & Gill, B. (2016). Is school value added indicative of principal quality? *Education Finance and Policy*, *11*, 283–309.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310.
- Collaborative for Academic, Social, and Emotional Learning. (2005). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs* (Illinois ed.). Chicago, IL: Author.
- Davidson, L., Crowder, M., Gordon, R., Domitrovich, C., Brown, R., & Hayes, B. (2018). A continuous improvement approach to social and emotional competency measurement. *Journal of Applied Developmental Psychology*, *55*, 93–106.

- Deming, D. (2014). Using school choice lotteries to test measures of school effectiveness. *American Economic Review: Papers and Proceedings*, *104*, 406–411.
- Durlak, J., Weissberg, R., Dymnicki, A., Taylor, R., & Schellinger, K. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432.
- Dweck, C. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
- Education Analytics. (2017). *Growth Model for Educator Evaluation 2016/17* (Technical report, prepared for the New York State Education Department by Education Analytics). Madison, WI.
- Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2016). Selecting growth measures for use in school evaluation systems: Should proportionality matter? *Educational Policy*, *30*, 465–500.
- Fuller, W. (1987). *Measurement error models*. Hoboken, NJ: John Wiley.
- Gawade, N., & Meyer, R. (2016). Measuring teacher effectiveness using value-added models of high school achievement. *Teachers College Record*, *118*, 1–32.
- Gehlbach, H., & Hough, H. (2018). *Measuring social emotional learning through student surveys in the CORE districts: A pragmatic approach to validity and reliability* (PACE Working paper). Retrieved from http://edpolicyinca.org/sites/default/files/SEL_Validity_May-2018.pdf
- Gershenson, S. (2016). Linking teacher quality, student attendance, and student achievement. *Education Finance and Policy*, *11*, 125–149.
- Grissom, J., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. *Educational Evaluation and Policy Analysis*, *37*, 3–28.
- Guarino, C., Reckase, M., & Wooldridge, J. (2015). Can value-added measures of teacher performance be trusted? *Education Finance and Policy*, *10*, 117–156.
- Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools*, *4*, 221–239.
- Hamre, B. K., & Pianta, R. C. (2006). Student–teacher relationships. In G. G. Bear & K. M. Minke (Eds.), *Children's needs III: Development, prevention, and intervention* (pp. 59–71). Washington, DC: National Association of School Psychologists.
- Heckman, J. J., & Rubinstein, Y. (2001). The importance of noncognitive skills: Lessons from the GED testing program. *American Economic Review*, *91*, 145–149.
- Jackson, K. (2018). What do test scores miss? The importance of teacher effects on non-test-score outcomes. *Journal of Political Economy*, *126*, 2072–2107.
- Jennings, J. L., & DiPrete, T. A. (2010). Teacher effects on social and behavioral skills in early elementary school. *Sociology of Education*, *83*, 135–159.
- Jennings, P. A., & Greenberg, M. T. (2009). The prosocial classroom: Teacher social and emotional competence in relation to student and classroom outcomes. *Review of Educational Research*, *79*, 491–525.
- Kane, T., & Staiger, D. (2001). *Improving school accountability measures* (National Bureau of Economic Research, working paper 8156). Cambridge, MA: National Bureau of Economic Research.
- Kane, T., & Staiger, D. (2002). Volatility in school test scores: Implications for test-based accountability systems. In D. Ravitch (Ed.), *Brookings papers on education policy* (pp. 235–283). Washington, DC: Brookings Institution.

School Differences in Social–Emotional Learning Gains

- Kautz, T., Heckman, J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success*. Paris, France: Organization for Economic Co-Operation and Development.
- Koedel, C., & Betts, J. (2010). Value added to what? How a ceiling in the testing instrument influences value-added estimation. *Education Finance and Policy*, 5, 54–81.
- Ladd, H. F., & Sorensen, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy*, 12, 241–279.
- Ladd, H. F., & Walsh, R. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21, 1–17.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. St. Paul, MN: Center for Applied Research and Educational Improvement.
- Liu, J., & Loeb, S. (2018). *Engaging teachers: Measuring the impact of teachers on student attendance in secondary school* (CEPA working paper No. 17-01). Retrieved from <http://cepa.stanford.edu/wp17-01>
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- McCaffrey, D., Lockwood, J. R., Koretz, D., Louis, T., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29, 67–101.
- 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, 101–119.
- Meyer, R. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review*, 16, 283–301.
- Meyer, R., & Dokumaci, E. (2015). Value-added models and the next generation of assessments. In R. W. Lissitz (Ed.), *Value added modeling and growth modeling with particular application to teacher and school effectiveness* (pp. 139–190). Charlotte, NC: Information Age.
- Meyer, R., Wang, Y., & Rice, A. (2018). *Measuring students' social-emotional learning among California's CORE districts: An IRT modeling approach* (PACE working paper). Retrieved from http://edpolicyinca.org/sites/default/files/Measuring_SEL_May-2018.pdf
- Muraki, E. (1992). *A generalized partial credit model: Application of an EM algorithm* (ETS research report 92-06). Princeton, NJ: Educational Testing Service.
- Olweus, D. (1994). Bullying at school. In L. R. Huesmann (Ed.), *Aggressive behavior: Current perspectives* (pp. 97–130). Boston, MA: Springer.
- Raudenbush, S., & Willms, D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20, 307–355.
- Reardon, S., & Raudenbush, R. (2009). Assumptions of value-added models for estimating school effects. *Education Finance and Policy*, 4, 492–519.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73, 417–458.
- Ruzek, E. A., Domina, T., Conley, A. M., Duncan, G. J., & Karabenick, S. A. (2015). Using value-added models to measure teacher effects on students' motivation and achievement. *The Journal of Early Adolescence*, 35, 852–882.

- SAS. (2018). SAS EVAAS: Statistical Models and Business Rules of OH EVAAS Analyses. Prepared for Ohio Department of Education. Cary, NC: SAS. Retrieved from <http://education.ohio.gov/getattachment/Topics/Data/Accountability-Resources/Value-Added-Technical-Reports-1/Technical-Documentation-of-EVAAS-Analysis.pdf.aspx>
- Tekwe, C., Carter, R., Ma, C. X., Algina, J., Lucas, M., & Roth, J., . . . Resnick, M. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29, 11–36.
- Thomsen, J. (2013). *50-state comparison: State school accountability report cards*. Retrieved from <https://www.ecs.org/state-school-accountability-report-cards/>
- Transforming Education. (2016). *Measuring MESH: Student and teacher surveys curated for the CORE districts*. Retrieved from https://www.transformingeducation.org/wp-content/uploads/2017/04/160406_MeasuringMESH_ForRelease2.pdf
- Ttofi, M. M., & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology*, 7, 27–56.
- Waters, T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Aurora, CO: Mid-Continent Research for Education and Learning.
- West, M. R., Buckley, K., Krachman, S. B., & Bookman, N. (2018). Development and implementation of student social-emotional surveys in the CORE districts. *Journal of Applied Developmental Psychology*, 55, 119–129.
- West, M. R., Pier, L., Fricke, H., Hough, H., Loeb, S., Meyer, R. H., & Rice, A. (2018). *Trends in student social-emotional learning: Evidence from the CORE Districts* (PACE working paper). Retrieved from https://edpolicyinca.org/sites/default/files/SEL_Trends_May-2018.pdf

Authors

SUSANNA LOEB is a professor of education and public affairs and the director of the Annenberg Institute at Brown University, 164 Angell Street, Providence, RI 02912, USA; email: susanna_loeb@brown.edu. She was formerly the Barnett Family Professor of Education and founding director of the Center for Education Policy Analysis at Stanford University. Her research interests are educational outcomes and their distribution; school finance and governance; and educator development, effectiveness, and labor markets.

MICHAEL S. CHRISTIAN is a research scientist at Education Analytics, 131 W. Wilson St., Suite 200, Madison, WI 53703, USA; email: mchristian@edanalytics.org. His research interests are student growth measurement, educator impact measurement, and school accountability models.

HEATHER HOUGH is the executive director of Policy Analysis for California Education, 520 Galvez Mall, CERAS Room 401, Stanford, CA 94305, USA; email: hjough@stanford.edu. Her research interests are district- and state-level policymaking and implementation, with a focus on policy coherence, system improvement, and school and teacher accountability.

School Differences in Social–Emotional Learning Gains

ROBERT H. MEYER is the president and CEO of Education Analytics, 131 W. Wilson St., Suite 200, Madison, WI 53703, USA; email: rhmeyer@edanalytics.org. His research interests are value-added modeling, evaluation methods, strategic management of human capital, and social–emotional learning.

ANDREW B. RICE is the chief operating officer at Education Analytics, 131 W. Wilson St., Suite 200, Madison, WI 53703, USA; email: abrace@edanalytics.org. His research interests are predictive modeling, school accountability models, assessment data, social–emotional learning, and data systems.

MARTIN R. WEST is a professor of education at the Harvard Graduate School of Education, 6 Appian Way Gutman 454, Cambridge, MA 02138, USA; email: martin_west@gse.harvard.edu. His research interests are the politics of K–12 education in the United States and how education policies affect student learning and noncognitive development.

Manuscript received April 4, 2018

First revision received November 5, 2018

Second revision received February 20, 2019

Accepted March 18, 2019