

Trends in Student Social-Emotional Learning: Evidence from the CORE Districts

Martin R. West

Harvard University

Libby Pier

Education Analytics

Hans Fricke

Policy Analysis for California Education

Heather Hough

Policy Analysis for California Education

Susanna Loeb

Stanford University

Robert H. Meyer

Education Analytics

Andrew B. Rice

Education Analytics

Mounting evidence demonstrates that social-emotional skills are important for students' academic and life success, yet we have limited evidence on how these skills develop over time and how this development varies across student subgroups. In this study, we use the first large-scale panel survey of social-emotional learning (SEL) to describe how four SEL constructs—growth-mindset, self-efficacy, self-management, and social awareness—develop from Grade 4 to Grade 12, and how these trends vary by gender, socioeconomic status, and race/ethnicity. Our results are based on self-report student surveys administered in the 2014-15 and 2015-16 school years to roughly 400,000 students within the CORE districts, a network of urban school districts in California. We simulate cohort trends for students continuously enrolled in six districts within CORE by adding the mean change at each grade level observed among students surveyed in both years to the mean scores of Grade 8 students in 2015-16. We find that, with the exception of growth mindset, social-emotional skills do not increase monotonically as students move through school. Self-efficacy, social awareness, and to a lesser degree self-management decrease after Grade 6. Trends in SEL also vary by subgroup. Female students report higher self-management and social awareness than boys, but their self-efficacy drops sharply relative to boys in middle and high school. Economically disadvantaged students report lower levels of each social-emotional construct, but gaps in self-management, growth mindset, and self-efficacy narrow in high school. White students report higher levels of social-emotional skills than African American and Latinx students; Asian students report levels of self-management similar to White students but exhibit declining self-efficacy over time. While our findings cannot be generalized beyond the California districts we study, the scope and scale of our data far exceeds anything in the extant literature.

VERSION: May 2018

Acknowledgements

This paper was produced as part of the CORE-PACE Research Partnership, which is focused on producing research that informs continuous improvement in the CORE districts (Fresno, Garden Grove, Long Beach, Los Angeles, Oakland, Sacramento City, San Francisco, and Santa Ana unified school districts) and policy and practice in California and beyond. We thank the leaders and administrators in the CORE districts for their support throughout this project, and the generous funder of this research, the Walton Family Foundation. PACE working papers are circulated for discussion and comment purposes and have not undergone the peer-review process that accompanies official PACE publications.

Student skills that are not captured by tests of academic achievement and ability predict a range of academic and life outcomes (Almlund, Duckworth, Heckman, & Kautz, 2011; Heckman, Humphries, & Kautz, 2014; Deming, 2017). Both intrapersonal skills (such as the ability to regulate one’s behavior in pursuit of long-term goals) and interpersonal skills (such as the ability to collaborate with others) appear to be key complements to cognitive ability in determining students’ success in school, post-secondary education, and the labor market (National Research Council, 2012). In addition, such “non-cognitive” or “social-emotional” skills may be more malleable in school settings than cognitive abilities, making them attractive targets for interventions aimed at improving student success (Cunha & Heckman, 2008; Dee & West, 2011; Heckman & Kautz, 2013). Consistent with this logic, a recent meta-analysis finds that school-wide interventions targeting social-emotional learning (SEL) generate improvements in students’ academic achievement (Durlak, Dymnicki, Taylor, Weissberg, & Schellinger, 2011).

Accumulating evidence on the importance of non-tested skills has led policymakers to look beyond test scores when seeking to measure and improve student outcomes. The recently enacted federal Every Student Succeeds Act (ESSA), for example, requires states to incorporate an “additional indicator of school quality or student success” not based on math and reading test scores into their school accountability systems. A growing number of states have established standards for SEL or incorporated social-emotional skills into their academic content standards (Dusenbury et al., 2015). Meanwhile, the Aspen Institute has launched a National Commission on Social, Emotional, and Academic Development with a mandate to “re- envision what constitutes success in our schools” and “explore how schools can fully integrate social, emotional, and academic development to support the whole student.”¹

At the forefront of this trend are the CORE districts, a network of large urban districts in California serving nearly one million students. These districts received a waiver from the U.S. Department of Education in 2013 to implement an alternative to the school accountability system then-mandated under the No Child Left Behind Act.² The CORE districts used this flexibility to develop a measurement system that includes survey-based measures of SEL and school culture and climate alongside traditional academic indicators. Although the obligation to use its SEL survey for school accountability was voided by the 2015 enactment of ESSA, the CORE districts continue to collect data on SEL to guide school policy and continuous improvement.

The CORE districts’ ongoing partnership provides a unique opportunity to inform efforts to promote SEL with evidence on how key social-emotional skills develop as students progress through American schools. Policymakers need to know how social-emotional skills typically vary across grade levels and subgroups in order to interpret aggregate data on SEL and determine

¹ <https://www.aspeninstitute.org/programs/national-commission-on-social-emotional-and-academic-development/>

² The CORE districts that implemented the waiver are Fresno, Long Beach, Los Angeles, Oakland, San Francisco, and Santa Ana unified school districts. Garden Grove and Sacramento City unified school districts are also part of the CORE network.

where interventions or supports are most needed. Similarly, educators need such information in order to interpret data on their own students and take appropriate action. In some cases, evidence of trends in SEL has already informed the design of interventions. For example, evidence that many students experience a decline in self-esteem and school engagement as they move from elementary school to middle school (Blum & Libbey, 2004; Eccles, Lord, & Midgley, 1991; Eccles, Wigfield, Midgley, & Reuman, 1993) has motivated the development of SEL-focused interventions aimed at supporting students' through this transition (see, e.g., Blackwell, Trzesniewski, and Dweck, 2007).

However, there is a lack of research examining how a broader set of social-emotional skills develop over time, particularly for different student subgroups. Existing studies with a longitudinal design tend to focus on the development of SEL only in early childhood or elementary school (e.g., Edossa, Schroders, Weinert, & Artelt, 2018; Rothbart, Posner, & Kieras, 2006) or consider only a single SEL construct (Ross & Tolan, 2017). Cross-sectional studies in turn do not shed light on how skills evolve over time (e.g., Ablard & Lipschultz, 1998; Chodhury et al., 2012). Many studies of SEL rely on small convenience samples of students within specific settings (e.g., Duckworth, Tsukayama, & May, 2010; Blackwell et al., 2007), raising questions about the generalizability of their findings. Moreover, variation in the specific constructs and measures used to assess students' social-emotional skills makes it difficult to compare results across studies (e.g., Berg et al., 2017; Duncan & Magnuson, 2011).

In this paper, we use the CORE districts' unique SEL survey to address two broad questions: How do the four SEL constructs assessed by that survey—growth mindset, self-efficacy, self-management, and social awareness—develop from Grade 4-12? And how do these patterns vary by gender, socioeconomic status, and race/ethnicity? Our analyses are based on self-report surveys administered to nearly 400,000 students in the 2014-15 and 2015-16 school years. With two years of data, we can only track the development of SEL for a given student over the course of a single school year. However, we are able to aggregate information on these changes across multiple grade levels in order to simulate long-term trends for students expected to remain enrolled in participating districts through middle and high school. More specifically, we calculate mean score gains for students who completed the survey in both years, and we use these gains to extrapolate from Grade 8 (the midpoint of our sample) to both prior and subsequent grades. The results of these simulations show how the SEL constructs develop among students who would be expected to attend schools in participating districts continuously from Grade 4 through Grade 12, assuming that everything else about those districts (including selection into and out of the districts, as well as all aspects of the educational environment relevant to SEL development) remain as they were in the 2015-16 school year.

In reporting these trends, we emphasize that the measures gathered by the CORE districts' SEL survey are self-reports and therefore reflect students' subjective assessments of their social-emotional skills. Students evaluating their own skills must employ an external frame of reference in order to reach a judgment about their relative standing. As a result, differences in self-reports over time or across students may reflect differences in normative standards rather than authentic differences in skills—a phenomenon known as *reference bias* (West et al.,

2016). Students' responses may also be influenced by *social desirability bias* (Paulhus, 1991), or the tendency of survey respondents to offer positive self-descriptions, and students' susceptibility to this bias also may vary across students and over time. Finally, certain students' responses may also be influenced by cultural differences that lead them to interpret or respond to items in different ways, or by differences in their home or school environments that influence their ability to demonstrate a given social-emotional skill. We show that students' self-reports of each SEL construct are associated in expected ways with theoretically related academic and behavioral indicators, providing at least partial evidence of validity. Even so, we urge caution when interpreting changes in these self-report measures over time and differences in both levels and trends across subgroups. Although the patterns in students' self-descriptions that we document are of interest in and of themselves, they do not necessarily capture true differences in underlying skills.

These analyses complement and extend a set of recent studies using the CORE districts' SEL survey to provide new insight into the measurement and development of social-emotional skills. West, Buckley, Krachman, & Bookman (2018) describe how the four social-emotional competencies were selected for assessment, explain the process for curating and piloting student surveys, and provide preliminary evidence of the measures' reliability and validity as an indicator of school quality. Meyer, Wang, and Rice (2018) examine the psychometric properties of the SEL measures, including consistency of measurement across grades and demographic groups. They also use Item Response Theory to develop scale scores for each construct that we rely on in this paper. Finally, Loeb et al. (2018) produce and evaluate school-by-grade estimates of students' growth in each SEL construct. The current paper builds upon this body of work by providing additional evidence of the reliability and validity of the CORE district SEL measures and by conducting the first analyses comparing SEL competencies across grades and subgroups. While our findings cannot be generalized beyond the California districts we study and the measures they employ, the scope and scale of our data far exceeds anything in the extant literature.

Literature Review

Despite heightened interest and activity on the part of policymakers, there remains a lack of consensus regarding how different aspects of students' SEL evolve over time. This is in part because most investigations of students' SEL focus on students in early elementary grades (Ross & Tolan, 2017). Nevertheless, the literature suggests that, unlike academic achievement, children's and adolescents' social and emotional development does not proceed linearly or monotonically over time (e.g., Pintrich & Zusho, 2002; Schunk & Meece, 2006; Schunk & Pajares, 2002; West et al., 2016). Although students are able to engage in more strategic or metacognitive thinking as they age (Steinberg, 2007, Wigfield, Byrnes, & Eccles, 2006), adolescence is also when students' social-emotional arousal is at its peak. These patterns suggest that specific social-emotional competencies may develop differently depending on the degree to which they require emotional regulation relative to cognitive control.

In this section, we discuss the available evidence on the development of the four constructs assessed by the CORE districts' SEL survey. We then turn to differences in SEL

trajectories over time for subgroups defined based on gender, socioeconomic status, and race/ethnicity. The four constructs are defined as follows:

- *Self-management*, also referred to as self-control or self-regulation, 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 (CASEL, 2005).
- *Growth mindset* is the belief that one's abilities can grow with effort. Students with a growth mindset believe that they can develop their skills through effort, practice, and perseverance. These students embrace challenges, see mistakes as opportunities to learn, and persist in the face of setbacks (Dweck, 2006).
- *Self-efficacy* is the belief in one's ability to succeed in achieving an outcome or reaching a goal. Self-efficacy reflects confidence in the ability to exert control over one's own motivation, behavior, and environment and allows students to become effective advocates for themselves (Bandura, 1997).
- *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 and supports (CASEL, 2005).

Development of SEL

Studies generally suggest that *self-management* (also referred to as *self-regulation* or *self-control*) declines during early adolescence (Duckworth et al., 2010; West et al., 2016). However, studies focusing on how self-management develops throughout adolescence are limited (Gestsdottir & Lerner, 2008) and the evidence is mixed.³ Some researchers have suggested that certain skills required for self-management—such as controlling attention, inhibiting responses, and self-monitoring progress—actually *increase* as students age, but other variables related to self-management—such as interest in school, motivation, and the changing classroom environment—dampen students' ability to exercise self-management (e.g., Gestsdottir & Lerner, 2008; Pintrich & Zusho, 2002). Therefore, although self-management may decline as students age, this likely depends on whether the measures used are also capturing related competencies or contextual variables.

In terms of *growth mindset*, few studies have examined students' trajectories over time; the literature has focused instead on the relationship of growth mindset and academic outcomes (e.g., Blackwell et al., 2007; Claro, Paunesku, & Dweck, 2016) and the effects of interventions seeking to foster a growth mindset (e.g., Dweck, 2006; Yeager, et al., 2016). Among studies examining changes in growth mindset, the findings are ambiguous. Some researchers have reported that growth mindset decreases during middle school (Pintrich &

³ In contrast, improvements in self-management through infancy and early childhood are well documented (e.g., Edossa et al., 2018; Kopp, 1982; Rothbart et al., 2006; Rueda, Posner, & Rothbart, 2005).

Zusho, 2002), whereas others show growth mindset may in fact increase during this period (West et al., 2016).

By contrast, a large body of work has established that *self-efficacy* tends to decline in middle school (e.g., Anderman, Maehr, & Midgley, 1999; Pajares & Valiante, 1999; Pintrich & Zusho, 2002; Schunk & Meece, 2006; Schunk & Pajares, 2002; Urda & Midgley, 2003; Wigfield et al., 2006). This may be because younger students tend to overestimate their capabilities and students' expectations become more realistic as they age (Pintrich & Zusho, 2002). Increased emphasis on competition, social comparison, and norm-referenced grading during the transition to middle school may heighten or enhance the accuracy of comparative self-assessment (Anderman et al., 1999; Urda & Midgley, 2003; Wigfield, Byrnes, & Eccles, 2006). Indeed, the middle school transition is a particularly vulnerable time for students' self-efficacy beliefs (Schunk & Pajares, 2002; Schunk & Meece, 2006). However, some studies suggest that domain-specific self-efficacy (i.e., ELA- or math-specific) increases during middle school (Shell, Colvin, & Bruning, 1995; Zimmerman & Martinez-Pons, 1990). This finding suggests that more global measures of self-efficacy may manifest different developmental patterns than domain-specific measures.

Finally, research into the development of *social awareness* indicates that students become more socially aware over time as peer groups become more central (Ryan, 2001; Rubin, Coplan, Chen, Buskirk, & Wojslawowicz, 2005; Wigfield, Byrnes, & Eccles, 2006). Some concepts or skills that are related to or prerequisites of social awareness—such as self-awareness, self-reflection, perspective taking, and metacognition—improve as students age (Choudhury et al., 2006; Eccles, 1999; Piaget, 1972; Yurgelun-Todd, 2007). However, because empirical studies of social awareness generally focus on social skills broadly, it is difficult to disentangle particular developmental patterns for social *awareness* specifically (Farrington et al., 2012). In addition, some studies find, contrary to the above results, that younger students respond more positively to measures of social skills (Gaspar, Cerquiera, Branquinho, & Gaspar de Matos, 2018).

Development of SEL by Subgroup

In addition to varying developmental trends for specific SEL constructs, research has suggested that the development of these constructs differs across student groups. Here we review trends by gender, socioeconomic status, and race/ethnicity.

Gender. There is good reason to believe that SEL trajectories differ according to students' gender, but the empirical evidence to date has been inconclusive. Early adolescence is a time when culturally-relevant gender stereotypes intensify; puberty renders these stereotypes more salient and can reshape teens' perceptions of themselves and others (Eccles, 1987; Hill & Lynch, 1983; Kågesten et al., 2016). Furthermore, differences between boys' and girls' biological development can manifest in their non-cognitive skills. For example, girls tend to display higher degrees of self-management than boys in elementary school and early adolescence (Ablard & Lipshultz, 1998; Duckworth & Seligman, 2006; Moffitt et al., 2011; Pintrich & Zusho, 2002; Zimmerman & Martinez-Pons, 1990).

Evidence of gender differences in self-efficacy is more ambiguous. Some studies suggest that boys and girls have similar self-efficacy in elementary school, but girls display lower self-efficacy during the transition to middle school (e.g., Anderman, et al., 1989; Wigfield, Eccles, Maclver, Reuman, & Midgley, 1991; Wigfield, Eccles, & Pintrich, 1996). Yet, other studies suggest that gender differences during this time interact with students' level of academic achievement and stereotypic gender beliefs (Eccles, Adler, & Meece, 1984; Schunk & Pajares, 2002). As a result, gender differences in self-efficacy may be domain-specific, with girls tending to show higher levels than boys of self-efficacy in ELA but lower levels in math (Eccles, Wigfield, & Schiefele, 1998; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Marsh, 1989; Schunk & Meece, 2006; Wigfield et al., 1991). In contrast, some studies find no evidence of gender differences in self-efficacy (e.g., Pajares, 1996; Pajares & Graham, 1999; Roeser, Midgley, & Urdan, 1996; Smith, Sinclair, & Chapman, 2002).

In terms of growth mindset, girls are generally viewed as being more likely than boys to endorse a fixed mindset, rather than a growth mindset (Dweck, 1986, 2000; Dweck & Simmons, 2014; Halvorson, 2011), particularly when asked about their abilities in stereotypically male domains such as math or science (i.e., male-stereotyped domains; Farrington et al., 2012). However, multiple studies have found no relationship between gender and growth mindset (e.g., Macnamara & Rupani, 2017; Storek & Furnham, 2013; Tucker-Drob, Briley, Engelhardt, Mann, & Harden, 2016).

Finally, research suggests that girls display higher social awareness than boys during the transition to middle and high school. Because boys and girls experience distinct socialization practices during adolescence (Kågesten et al., 2016), girls' relationships to peers and families might differentially affect how they develop self- and social awareness. Specifically, Wentzel (1994) found that girls tend to display higher degrees of social behavior, social acceptance, social goal pursuit, and perceived social support. Gaspar and colleagues (2018) showed that although boys respond more positively to measures of subjective well-being, social support, problem solving, and emotional regulation, girls responded more positively to measures of basic social skills and interpersonal relationships. Similarly, studies have indicated that girls tend to suffer from intrapersonal behavior challenges, whereas boys tend to suffer from interpersonal behavior challenges (Hatzchristou & Hopf, 1996; Underwood, 2004). This discrepancy can be further exacerbated by teachers' differential expectations of boys' and girls' social behavior (Sonja, Milena, Jana, & Cirila, 2009).

Socioeconomic status. Relatively few studies examine how students' socioeconomic status (SES) is associated with social-emotional competencies (Schunk & Meece, 2006), but the bulk of the available evidence suggests the existence of gaps in SEL favoring economically advantaged students. For example, evidence suggests that students from economically disadvantaged backgrounds lag behind their peers in skills related to self-management, such as emotional regulation (Papini, Farmer, Clark, Micka, & Barnett, 1990), adaptability (Davis, 2012), or impulsive behavior (Takeuchi, Williams, & Adair, 1991). Factors related to self-efficacy have also been shown to vary by SES. For example, economically disadvantaged students may have lower self-esteem (Bolger, Patterson, Thompson, & Kupersmidt, 1995), they may be more likely to experience learning challenges early in school that dampen their self-efficacy later on

(Schunk & Miller, 2002), and their parents may have reduced expectations of their academic success (Alexander & Entwisle, 1998). In terms of growth mindset, a recent study found that students in Grades 4-12 attending schools with a higher concentration of students in poverty reported lower levels of growth mindset (Snipes & Tran, 2017). Finally, research shows that students from economically disadvantaged backgrounds may lag in competencies related to social awareness, as they may be more likely to struggle with peer relationships (Bolger, et al., 1995) or social competence (Winer & Thompson, 2013).

More generally, a large body of evidence shows that growing up in poverty is a major risk factor for low levels of social and emotional well-being in adolescence and adulthood (Bradley & Corwyn, 2002; Brooks-Gunn & Duncan, 1997; Kupersmidt, Griesler, DeRosier, & Patterson, 1995; Takeuchi et al., 1991; Yoshikawa, Aber, & Beardslee, 2012). However, much of this research focuses on behavioral indicators, such as aggression (e.g., Colder, Mott, Levy, & Flay, 2000; Sinclair, Pettit, Harrist, Dodge, & Bates, 1994) or misconduct (e.g., McCoy, Frick, Loney, & Ellis, 1999); on psychological well-being and mental illness (e.g., McLeod & Shanahan, 1993; Ortega & Corzine, 1990); or on physical well-being, such as health (e.g., Starfield, 1989) or substance abuse (e.g., Wills, McNamara, & Vaccaro, 1995). Evidence on the role of SES also may be sensitive to the measurement of SEL. Many studies showing a link between SES and SEL rely on parent or teacher reports of student behaviors (e.g., Duncan, Brooks-Gunn, & Klebanov, 1994, McLeod & Shanahan 1993; McLoyd, 1998), whereas one study of adolescent *self*-reports did not find a relationship between poverty and SEL (Conger, Conger, & Elder, 1997).

Race/ethnicity. Most of the literature examining racial and ethnic subgroup differences in SEL consist of cross-sectional studies measuring SEL at a single time point. Few studies have examined racial or ethnic differences in how adolescents develop self-regulation skills related to self-management (Pintrich & Zusho, 2002). For self-efficacy, there are mixed results. Some studies show no difference across different racial/ethnic groups (e.g., Britner & Pajares, 2001; Roeser et al. 1996). Other studies find that Asian students report lower self-efficacy (Eaton & Dembo, 1997), that Latinx teens report lower self-efficacy in writing (Pajares & Johnson, 1996), and that African American teens report lower self-efficacy in math (Pajares & Kranzler, 1995); still other studies suggest that certain underrepresented minority students may report *higher* degrees of their sense of academic competence (e.g., Graham, 1994). Few studies explicitly examine differences in growth mindset among adolescents of differing racial or ethnic background. One recent report found that African American and Latinx students self-reported lower levels of growth mindset than their White counterparts, but this finding was based on evidence from a single school district (Snipes & Tran, 2017). Finally, evidence of students' social awareness based on race or ethnicity is limited to elementary and middle school, and findings are ambiguous. DiPerna and Elliott (1999) describe how minority elementary school students were rated lower than their White counterparts on teacher-reported measures of interpersonal skills, but Malecki and Elliott (2002) report no differences in teachers' reports of social skills for White and minority students. Additionally, Wentzel (1994) showed that White middle schoolers were rated as more socially adept by their peers and teachers than African American students. So far, students' self-reported levels of social awareness have not been examined.

Summary

Taken together, an extensive body of research examining students' social and emotional skills at different points in time suggests that the development of students' social and emotional learning is construct-specific and may vary across different groups of students. Overall, the literature suggests self-management and self-efficacy may decrease in adolescence, while social awareness is expected to increase. Meanwhile, conflicting results regarding specific social skills highlight the need for more research examining trends in social awareness, and it is unclear whether growth mindset is expected to increase or decrease during elementary, middle, and high school. In terms of gender differences, research implies that girls are expected to have superior self-management and social awareness skills relative to boys, whereas girls are thought to have more of a fixed mindset and lower self-efficacy than boys—particularly with regard to male-stereotyped domains such as math and science. For differences by socioeconomic status, research indicates that students from economically disadvantaged backgrounds have lower SEL competencies relative to their more advantaged peers. However, since only a small proportion of the existing research focuses on SES differences, there is additional need for evidence supporting this hypothesis. As for racial and ethnic differences, there is little consensus as to whether minority students have higher or lower self-management, social awareness, or self-efficacy, and research surrounding minority students' growth mindset has mostly focused on the effectiveness of interventions related to growth mindset for improving minority students' academic achievement. It is possible that these conflicting findings stem from potential confounding with socioeconomic status, as well as with the specificity versus generality of the measures used to assess self-efficacy. The current study aims to shed additional light on to how these aspects of SEL develop over time as students progress through school.

Data and Methods

The CORE districts' SEL survey comprises a battery of items designed to measure four SEL constructs: self-management (9 items), social awareness (8 items), growth mindset (4 items), and self-efficacy (4 items). Students in Grades 4-12 rate themselves on the same 25 questions using a 5-point Likert scale. The same set of 25 questions was used in the 2014-15 and 2015-16 school years.

Measuring SEL development using these data requires us to transform the responses to the SEL items on the student survey into a metric. We create 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. Following Meyer, Wang, & Rice (2018), we use a generalized partial credit model (GPCM) to convert students' responses to these items into a scale score for each of the four constructs. Based on Muraki's (1992) extension of the partial credit model (PCM; Masters, 1982), GPCM can incorporate measures for which responses are on a multipoint scale, rather than only dichotomous items. The GPCM assigns more weight to items that better distinguish among students with different construct-specific abilities and appropriately accounts for missing student survey responses. Using a PCM in place of a GPCM to produce SEL scale scores yielded very similar substantive results, however, as did using raw scores.

Analytic Sample

Six CORE districts participated in the SEL survey in the 2014-15 and 2015-16 school years. These districts collectively serve roughly 572,000 students in Grades 4-12 across 1,200 schools. Approximately 390,000 (about 70%) of students in the districts completed the survey each year. Our analysis of the reliability and validity of the SEL measures is based primarily on the 2015-16 survey administration, though we also use data from the 2014-15 survey to examine the across-year reliability of the measures for students surveyed in both years. Our analysis of trends in the development of SEL across grades is also based on students surveyed in both years in order to address the possibility of non-random entry into and exit from schools within the CORE districts across grade levels.

As in any survey, not all students completed all items on the SEL survey. On average, each item was answered by 97.1% of the students across all grades in 2014-15 and by 97.5% of the students across all grades in 2015-16. Although the response rate on individual items was high, fewer than 70% of students completed all 25 survey items. Our final analytic sample included students who completed at least 50% of the items within each SEL construct. For example, the self-management construct has nine items in total. While analyzing self-management, we limited our sample to students who answered five or more self-management items. If a student answered five self-management items but skipped all growth mindset items, this student was included in analyses related to self-management but was excluded in analyses related to growth mindset. This analytic sample excluded only 0.47% of students from the full survey sample on average (across grades and constructs) in 2014-15 and 0.38% of students on average in 2015-16. Scale scores from the GPCM were used for all analyses to account for the remaining missing items.

Table 1 provides descriptive statistics for the analytic sample for the self-management construct in the 2015-16 school year (Panel A), as well as for all students enrolled in Grades 4-12 in the districts administering the survey that year (Panel B). Although the precise analytic sample varies across constructs because the inclusion criteria are applied separately to each construct, differences in the demographic composition of these samples are trivial. Students attending schools in the CORE districts in these grades are predominately Latinx (69 percent) and economically disadvantaged (73 percent); 36 percent are classified as English language learners. Sixty-eight percent of CORE students (389,211) were surveyed and responded to at least half of the nine items used to measure self-management. Relative to the full sample, students in the self-management analytic sample were roughly 3 percentage points more likely to be economically disadvantaged, 2 percentage points less likely to be African American, and 1 percentage point more likely to be Latinx, female, and not to have a disability. All other differences in demographic characteristics were smaller than a full percentage point. These patterns are generally consistent across grade levels.

Table 1. Descriptive Statistics for 2015-16 Sample

Panel A: Self-Management Analytic Sample

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Total
Male	50.9%	50.4%	50.4%	50.9%	50.5%	50.8%	51.0%	49.7%	49.7%	50.5%
Asian	6.8%	7.2%	7.9%	8.4%	7.9%	7.7%	8.3%	8.8%	9.3%	7.9%
African American	7.8%	7.7%	7.8%	7.7%	7.8%	8.1%	8.2%	8.4%	8.0%	7.9%
Latinx	70.9%	70.3%	69.1%	68.3%	69.3%	71.1%	70.4%	69.2%	68.7%	69.8%
ELL	50.4%	51.1%	42.6%	34.8%	32.0%	27.7%	24.6%	19.8%	18.8%	35.5%
SWD	11.5%	12.4%	12.0%	11.9%	11.2%	11.0%	10.0%	9.7%	11.1%	11.3%
Econ Dis	77.5%	76.7%	76.9%	76.4%	75.8%	76.5%	74.9%	74.0%	75.3%	76.1%
Foster	0.8%	0.8%	0.7%	0.6%	0.6%	0.6%	0.6%	0.5%	0.5%	0.7%
Homeless	2.6%	3.4%	2.5%	2.9%	3.2%	2.5%	2.8%	2.4%	2.3%	2.8%
N (students)	56,141	55,488	44,138	42,995	43,039	40,820	39,824	34,516	32,250	389,211
N (schools)	724	840	388	232	275	273	248	237	232	1,114

Panel B: All Students Enrolled in CORE Districts

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Total
Male	51.4%	51.1%	51.4%	51.4%	51.4%	52.1%	52.3%	50.8%	51.4%	51.5%
Asian	7.1%	7.3%	7.3%	7.4%	7.1%	6.5%	7.3%	7.8%	8.3%	7.3%
African American	9.4%	9.2%	9.4%	9.4%	9.7%	9.6%	10.1%	10.6%	10.5%	9.7%
Latinx	68.2%	68.0%	68.0%	67.8%	68.5%	70.2%	70.0%	68.5%	68.2%	68.6%
ELL	48.8%	50.1%	43.5%	36.2%	33.2%	30.9%	26.7%	22.2%	22.4%	35.5%
SWD	12.6%	13.4%	13.0%	12.9%	12.4%	11.6%	11.2%	10.6%	13.3%	12.4%
Econ Dis	74.2%	73.9%	74.9%	74.3%	73.7%	71.8%	70.6%	70.2%	70.9%	72.8%
Foster	0.9%	0.8%	0.7%	0.7%	0.7%	0.7%	0.8%	0.6%	0.6%	0.7%
Homeless	2.7%	3.3%	3.1%	3.3%	3.4%	2.9%	3.1%	2.9%	2.6%	3.0%
N (students)	72,213	69,326	62,495	61,170	60,984	67,716	64,622	56,071	58,375	572,972
N (schools)	787	912	436	269	328	313	278	274	271	1,213

Methods for Assessing Reliability and Validity

Prior to examining trends in SEL, we first employ several approaches to assess the reliability and validity of the measures gathered via the CORE districts SEL survey. For *reliability*, we examine the distributions of the SEL scale scores; we assess internal consistency using Cronbach’s coefficient alpha, including alpha normalized to the number of items in the longest scale (i.e., self-management, with nine items) using the Spearman-Brown formula (Brown, 1910; Spearman, 1910); and we compute across-year correlations for each construct to examine temporal stability. For *validity*, we examine inter-correlations among the four SEL measures by grade and each measure’s relationships with theoretically related academic and behavioral indicators. Specifically, we compute correlations between each SEL measure and

math and English language arts test scores and compare the mean number of days absent and days suspended for students within the bottom, middle, and top tercile of each SEL measure within their grade level.

Method for Simulating Trends

Because the data at hand are cross-sectional, we cannot interpret the means of each SEL construct by grade as observed in either the 2014-15 or 2015-16 school years as depicting true trends in students' SEL over time for two reasons. First, there could be idiosyncratic differences in SEL across grade cohorts. Second, there could be (and likely is) non-random entry into or exit out of the CORE districts across grade levels that is related to SEL. That is, students with particularly high or low levels of SEL may be systematically more likely to enter or exit the districts at specific grade levels. The method described below exploits the availability of repeated cross-sections from two consecutive years to address these issues.

In brief, we take the year-to-year changes in each construct for students who complete the survey for two consecutive years, and we anchor those changes to a specific mean (for Grade 8, which is the midpoint of our data) to produce “simulated cohort trends” and their associated standard errors. These simulated cohort trends can be interpreted as showing trends in the SEL constructs among students who would be expected to attend CORE district schools continuously from Grade 4 through Grade 12, assuming that everything else about the CORE districts (including selection into and out of the districts, as well as all aspects of the educational environment relevant to SEL development) remain as they were in 2014-15 and 2015-16. In other words, they represent a steady-state approximation of trends over time. This estimation method can equivalently be viewed as a model of student-level data that includes year effects and student-level effects.

For all analyses, we use a version of the GPCM true scores that have been standardized so that the mean score across grades within a construct is 0 and the standard deviation of the scores across grades within a construct is 1. The SEL scores were rescaled to this standardized scale to simplify presentation of the results, but the rescaling has no effect on the interpretation of the results.

To implement this approach, we first calculate the mean gain of SEL scores, $\delta^{\wedge}_{g(m)}$, for students who are enrolled in both years, which we refer to as the “matched sample” (i.e., 2015-16 standardized true score minus 2014-15 standardized true score) in each grade g in 2015-16 and subgroup m . For example, for students in Grade 9 in 2015-16, the gain is calculated by subtracting Grade 8 scores in 2014-15 from Grade 9 scores of the same students in 2015-16.

To obtain the simulated mean score in a given grade, $\mu^*_{g(m)}$, the mean gains (positive or negative) are added to the full sample mean score in a base grade, $\mu^{\wedge}_{g(m)}$. We choose Grade 8 as the base grade for the results presented in this paper, but this has no effect on the shape of the trend data; the choice of the base grade simply anchors the trend data around an actual cohort of students. Thus, the simulated score for each subgroup in grades after Grade 8 is as

follows (note that the summation of the delta term begins at 9, because it is the term for the 8-9 transition):

$$\mu_{g(m)}^* = \mu_{g(m)} + \sum_{h=9}^{h=g} \delta_{h(m)}$$

For grades earlier than 8 (i.e., Grade 7 and below), the following formula estimates the simulated score for each grade and subgroup (note that the summation of the delta term continues through Grade 8, because that term is for the 7-8 transition):

$$\mu_{g(m)}^* = \mu_{g(m)} - \sum_{h=g}^{h=8} \delta_{h(m)}$$

For example, for the simulated score in Grade 11, the gains of Grade 9, 10, and 11 are added to mean score in Grade 8. For the mean score in Grade 6, the gains of Grade 8 and 7 are subtracted from the mean score in Grade 8.

The error variances of $\mu_{g(m)}$ and $\delta_{g(m)}$ are given by the standard formulas for the variance of a mean. The error variance of the simulated score is equal to the sum of the error variances of the components. We assume that the covariance between the values of $\delta_{g(m)}$ is 0 since the means are based on different samples of students. The following formula shows this computation when simulating from Grade 8 to subsequent grades. The square root of this term produces the standard error.

$$\sigma_{g(m)}^2 = V(\mu_{g(m)}^*) = V(\mu_{g(m)}) + \sum_{h=9}^{h=g} V(\delta_{g(m)})$$

For grades earlier than Grade 8, the equivalent variance is calculated as follows:

$$\sigma_{g(m)}^2 = V(\mu_{g(m)}^*) = V(\mu_{g(m)}) + \sum_{h=g}^{h=8} V(\delta_{g(m)})$$

This method of simulating trends in SEL across grades requires us to focus on the subset of students who participated (and completed a majority of the items measuring a given construct) in both the 2014-15 and 2015-16 survey administrations. Appendix Table A1 provides descriptive statistics for this matched sample for the case of the self-management construct. The demographic characteristics of the students in this matched sample again correspond closely to those of the analytic sample for 2015-16 and of all students enrolled in CORE districts.

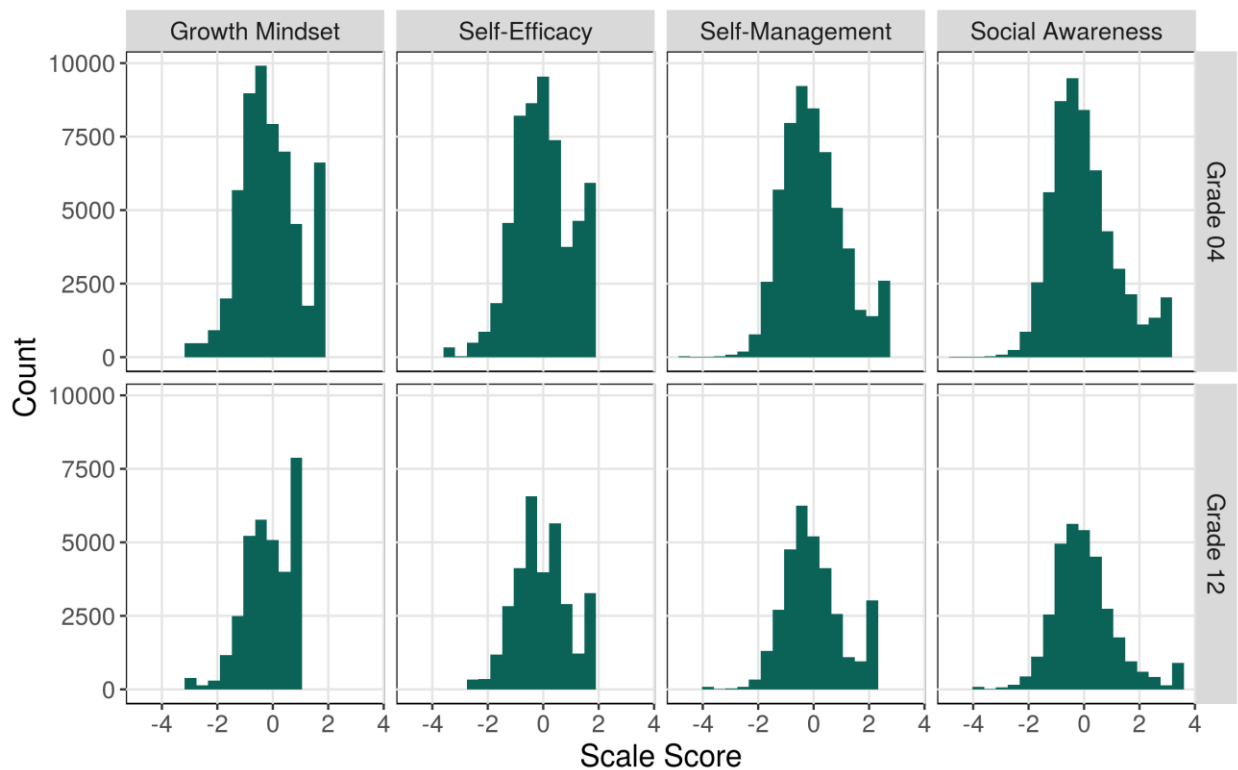
Results

We begin the presentation of our results by providing evidence on the reliability and the validity of the CORE districts' SEL measures. We then describe the results of the simulated trend analysis.

Evidence of Reliability

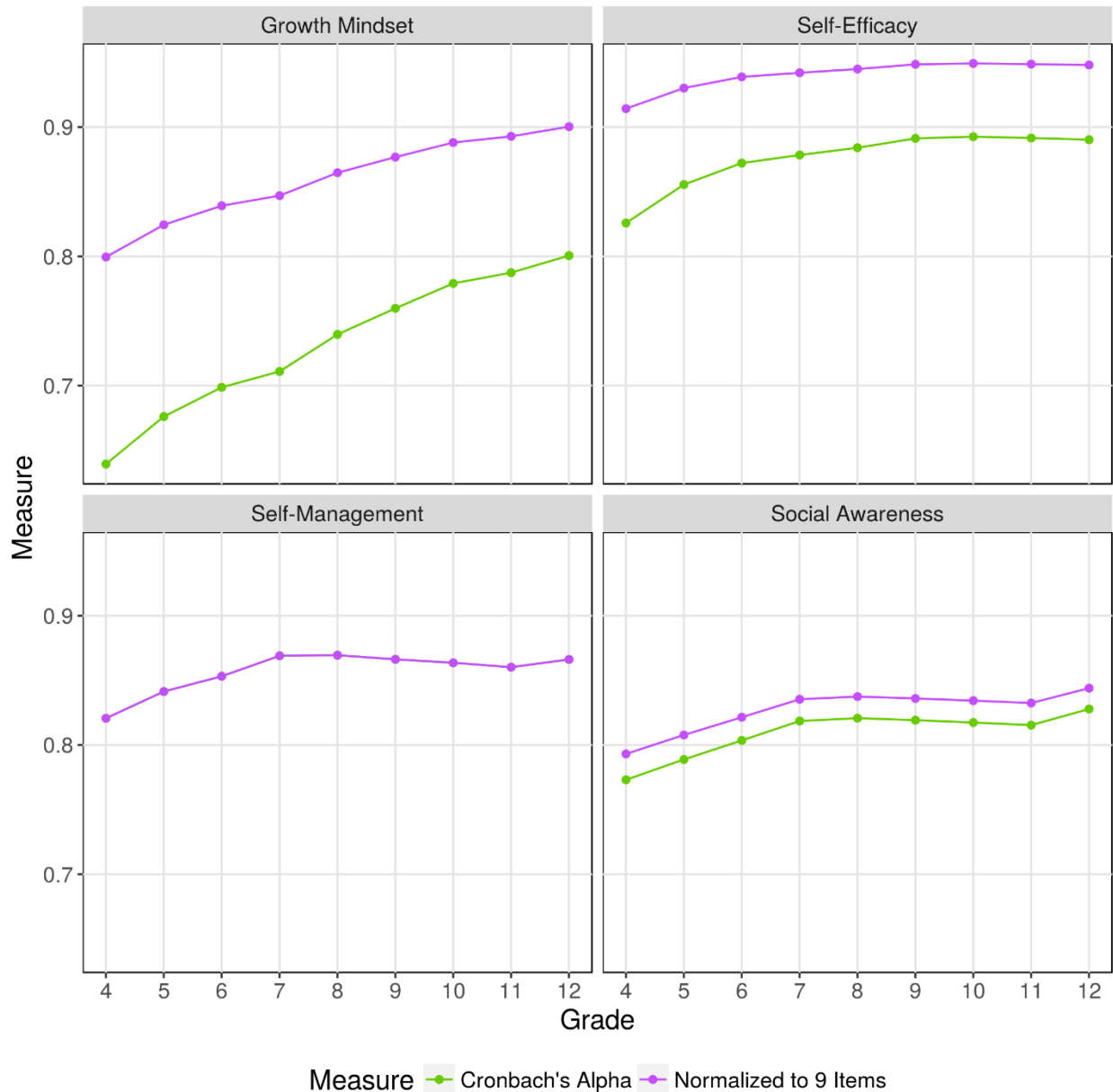
We first examine the distributions of student scale scores for each SEL construct. We find that the distributions of the SEL scale scores exhibit evidence of ceiling effects, with a substantial proportion of students choosing the most positive response to every item within a construct. Figure 1 illustrates this pattern by presenting histograms of the SEL scale scores in Grade 4 and Grade 12; histograms for other grades are similar. A ceiling effect is evident for all four constructs, but it is especially pronounced for the scale measuring growth mindset. These ceiling effects will present a challenge for efforts to analyze changes in SEL over time for individual students within the CORE districts. At the same time, they do not necessarily limit our ability to document trends across the full sample of surveyed students or numerically sizable subgroups. Each measure exhibits ample variation at both grade levels and, apart from the existence of ceiling effects, a roughly normal distribution.

Figure 1. Histogram of Distribution of SEL Scale Scores, Grades 4 and 12



We next assess the internal consistency of student self-reports by examining the inter-item correlation of the survey scales. Figure 2 depicts Cronbach's alpha for each scale, as well as alpha when normalized to nine items (i.e., the number of items in the self-management construct); because internal consistency decreases mechanically with the number of items, the normalized measure is more comparable across survey scales of varying length. We find that the measures generally demonstrate a high degree of internal consistency across all grade levels. The lone exception is the scale used to measure growth mindset, for which Cronbach's alpha drops below the commonly accepted benchmark of 0.7, particularly for younger students (e.g., fourth and fifth graders).

Figure 2. Internal Consistency of Social-Emotional Measures



In addition to internal consistency as measured by Cronbach’s alpha, another important aspect of reliability is temporal stability. In an effort to minimize disruption to instructional time, CORE elected not to administer its SEL measures to the same students multiple times the same year in a manner that would make it possible to calculate two-week test-retest reliabilities. We can, however, provide evidence on the measures’ temporal stability over one year among students who participated in both the 2014-15 and 2015-2016 surveys (Figure 3).

Figure 3. Across-Year Correlation of Social-Emotional and Test-Score Measures, 2014-15 and 2015-16

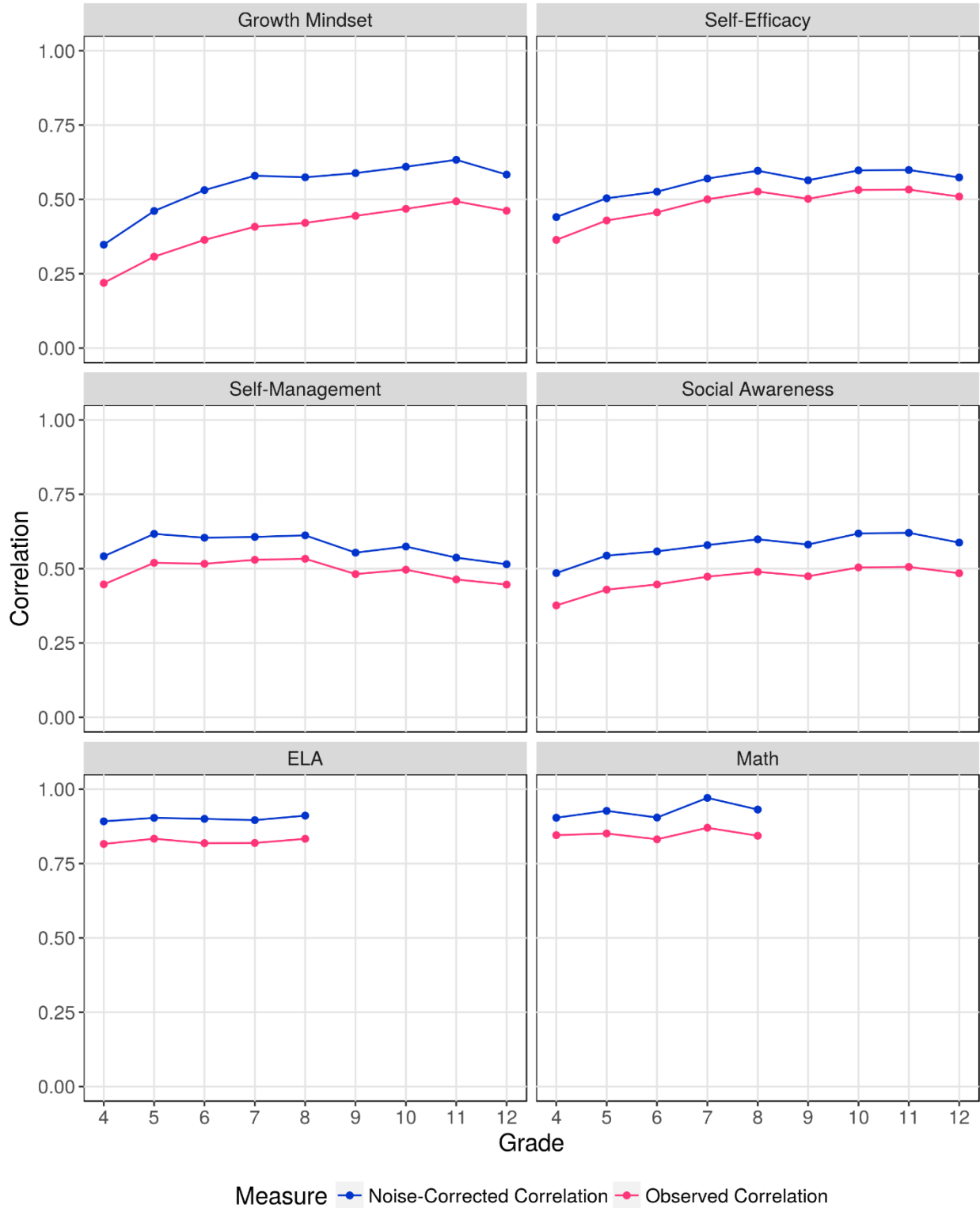


Figure 3 reports Pearson correlation coefficients for each SEL measure and, for purposes of comparison, state test scores through Grade 8. Noise-corrected correlations, which use our estimates of the measures' internal consistency to adjust for measurement error, are also displayed. As expected, students' math and English language arts (ELA) test scores are highly correlated from one year to the next ($r = 0.82 - 0.87$, $p < .001$), and the strength of this relationship is similar across grade levels. Although also statistically significant ($p < .001$), the parallel reliability estimates for the SEL measures are markedly lower ($r = 0.22 - 0.53$) and, with the exception of self-management, tend to increase across grades. The figure also confirms that the lower internal consistency of the SEL survey scales accounts for only a small fraction of the differences in stability between the SEL measures and test scores. Although striking, the lower temporal stability observed for the SEL measures is not necessarily a concern, given that one reason for educators' interest in SEL is the notion that they may be more malleable over time than cognitive ability.

Evidence of Validity

The four SEL measures are positively correlated with one another, though the strength of these relationships varies across constructs. Figure 4 plots these inter-correlations by construct and grade level. The correlations are consistently lowest for growth mindset. Meanwhile, the strongest relationships are observed for self-management and social awareness, with correlation coefficients between 0.5 and 0.6 depending on the grade level.

Each of the SEL measures is also related to the limited set of academic and behavioral indicators available in each district's administrative data: test scores, absences, and suspensions. Figure 5 plots the correlations between (i) each SEL measure and (ii) test scores in ELA and math in Grades 4-8 and 11—i.e., the grades in which California administers its state test (the Smarter Balanced Assessment Consortium). Two clear patterns are evident. First, the four SEL measures differ in the strength of their correlation with academic achievement as measured by state tests. The strongest relationships are generally observed for growth mindset and self-management, with correlations ranging between 0.3 and 0.4 between Grades 4 and 8. Interestingly, despite the fact that self-management and social awareness are strongly correlated with one another, they differ markedly in their relationship to test scores, with the correlations for social awareness hovering around 0.2 between Grades 4 and 8. Second, the relationship between each SEL construct and test scores in both subjects falls sharply between Grade 8 and 11.

The SEL measures are also predictive of student absences and suspensions. Students who assign themselves lower ratings on each construct miss more days and experience more out-of-school suspensions. Figure 6a plots the mean number of days absent for students in the bottom, middle, and top third within their grade level on each SEL measure; Figure 6b presents analogous information for the number of suspensions. We compare terciles of students when examining these outcomes because their non-normal distribution, with many students experiencing relatively few days absent and suspensions, renders correlation coefficients less informative. Other means of dividing students into groups (e.g., quartiles) yield comparable results.

Figure 4. Within-Year Correlations Among Social-Emotional Measures, 2015-16

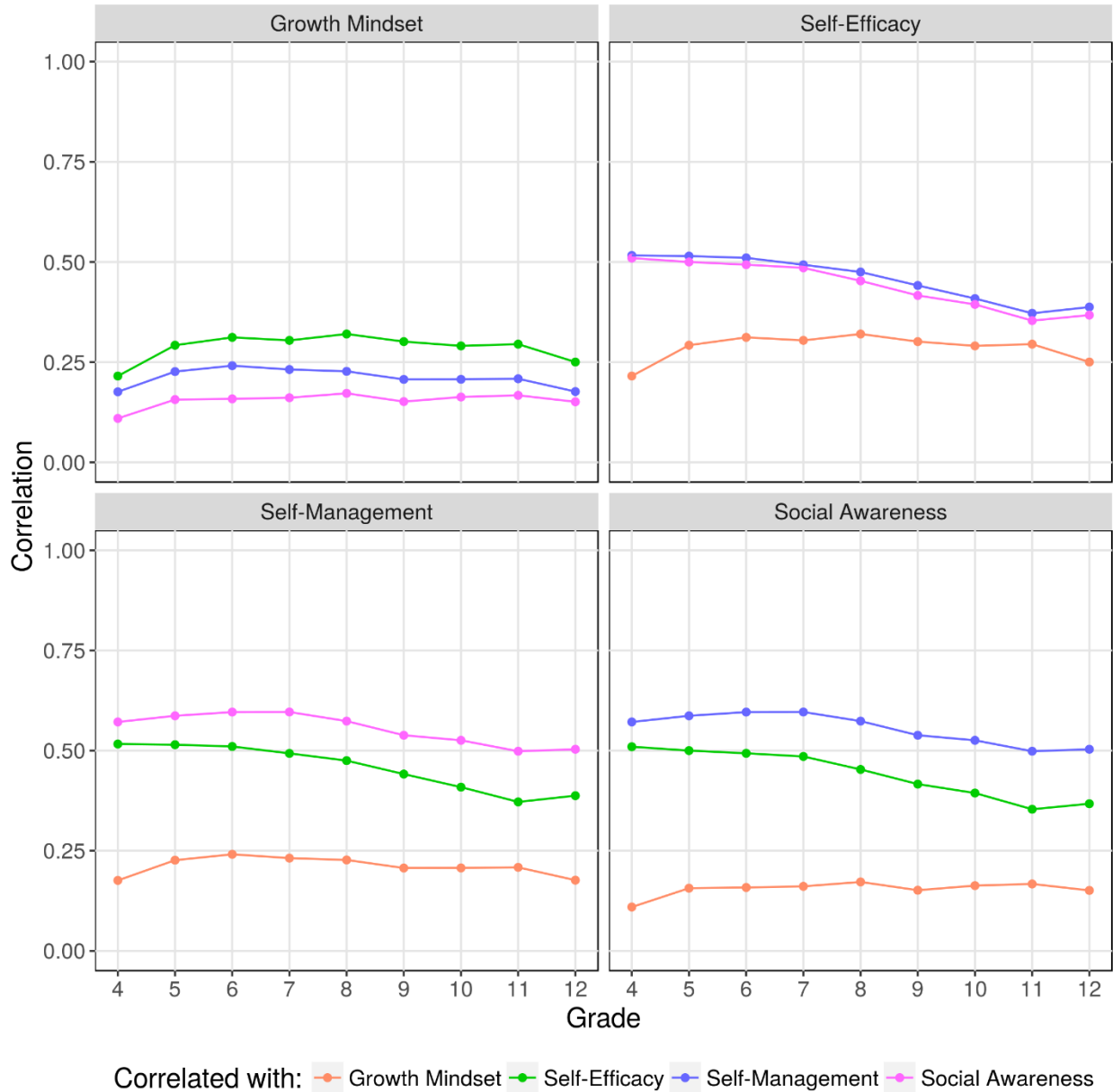


Figure 5. Within-Year Correlations Between Social-Emotional Measures and Test Scores, 2015-16

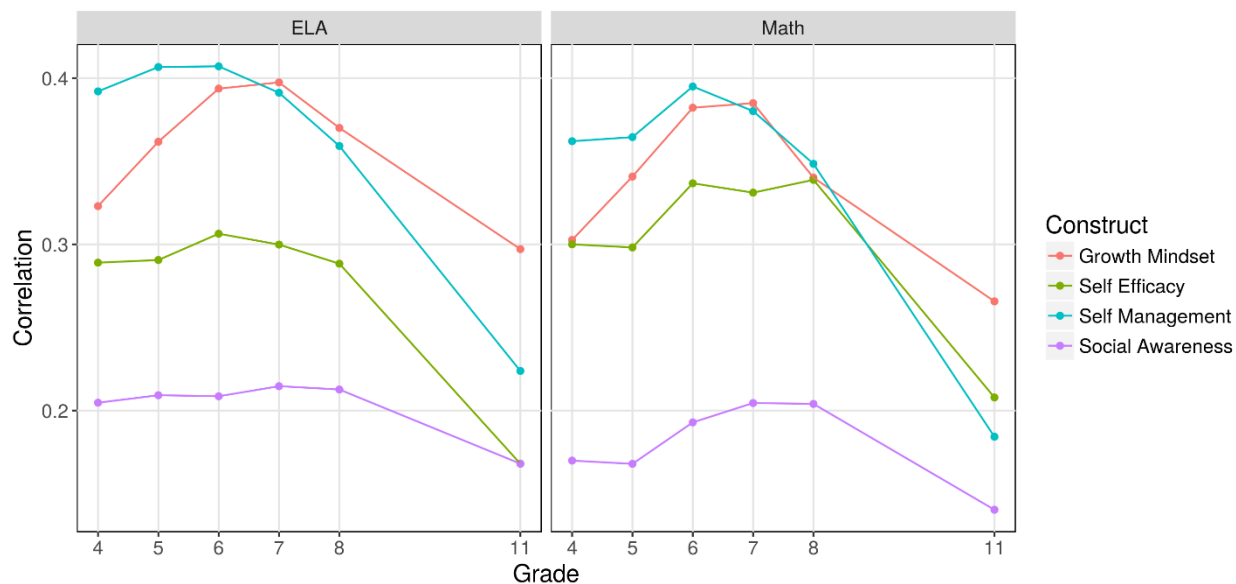


Figure 6. Mean Number of Days Absent and Suspensions by Social-Emotional Tercile, 2015-16

Panel a: Mean Number of Days Absent

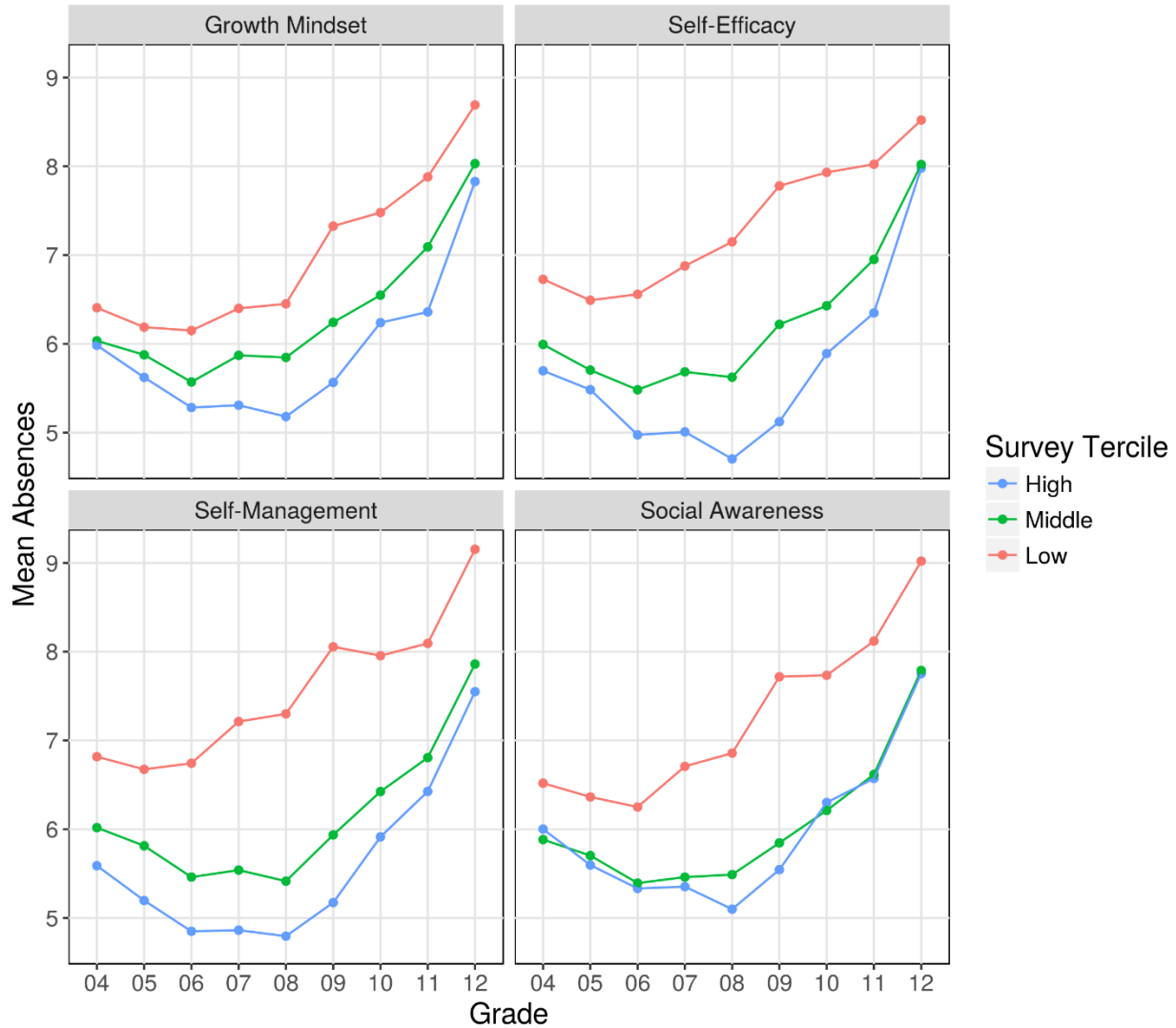


Figure 6a reveals that, at each grade level, students in the bottom third of each SEL measure are absent considerably more days, on average, than students in the top third; these gaps tend to expand between Grades 4 and 9, perhaps because students come to exercise greater autonomy in their absence behavior, before narrowing in the final years of high school. The gaps are largest for self-management, where a ninth-grade student in the bottom third is expected to be absent 2.9 additional days, or 56 percent more frequently, than a ninth-grade student in the top third. Similarly large gaps across the three groups are evident for self-efficacy, whereas gaps for social awareness and growth mindset are modestly smaller.

Panel b: Mean Number of Suspensions

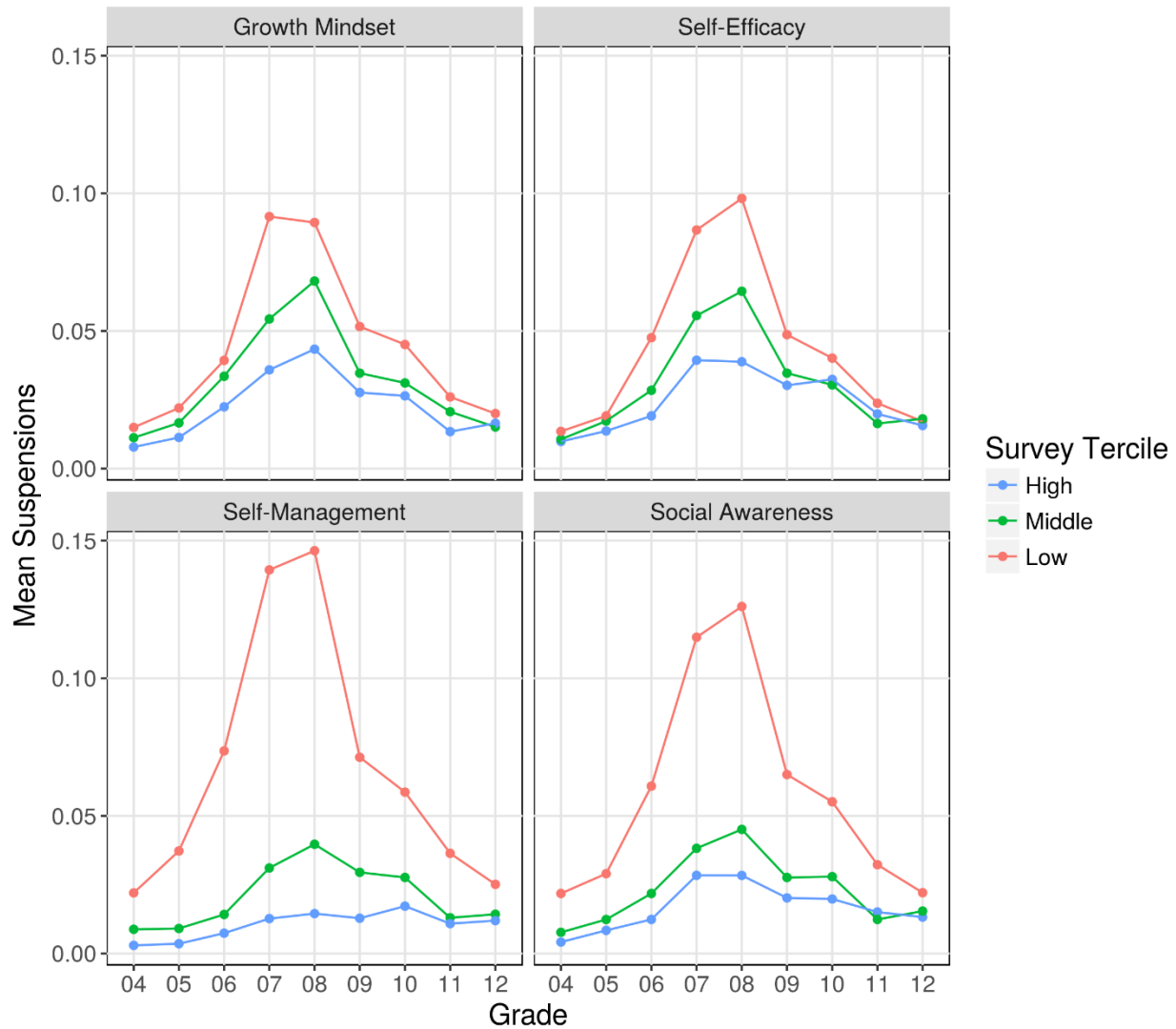


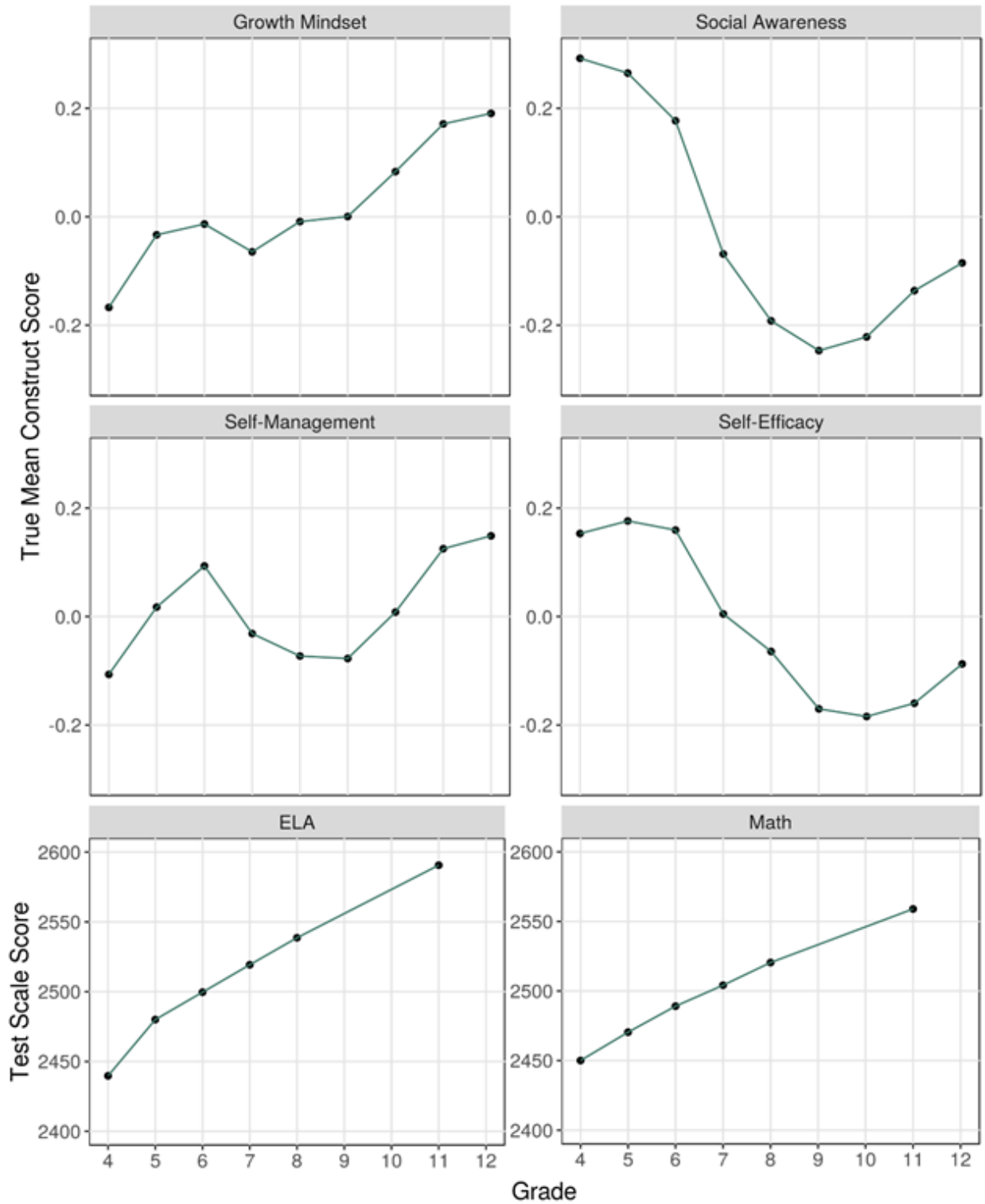
Figure 6b displays the parallel analysis for the number of suspensions, which peak in frequency in the CORE districts in Grades 7 and 8. We again see clear relationships for each SEL measure: Students who rate themselves critically are suspended more often. This relationship is most pronounced for self-management, where students in the bottom third experience 0.15 suspensions on average, more than seven times as many as students in the top third. Suspensions are much less frequent overall, and the gaps between students reporting different levels of SEL, in Grades 4-5 and Grades 11-12.

Trends in SEL Measures Across Grades

We now turn to the main results of this paper: simulated cohort trends of each SEL construct across grades for all students and various subgroups. In Figure 7, we first present the cross-sectional (i.e., non-simulated) mean scale scores for each SEL measure separately by grade level for Grades 4-12; ELA and math test scores are included for comparison purposes. It is immediately evident that, unlike the academic knowledge and skills captured by state test

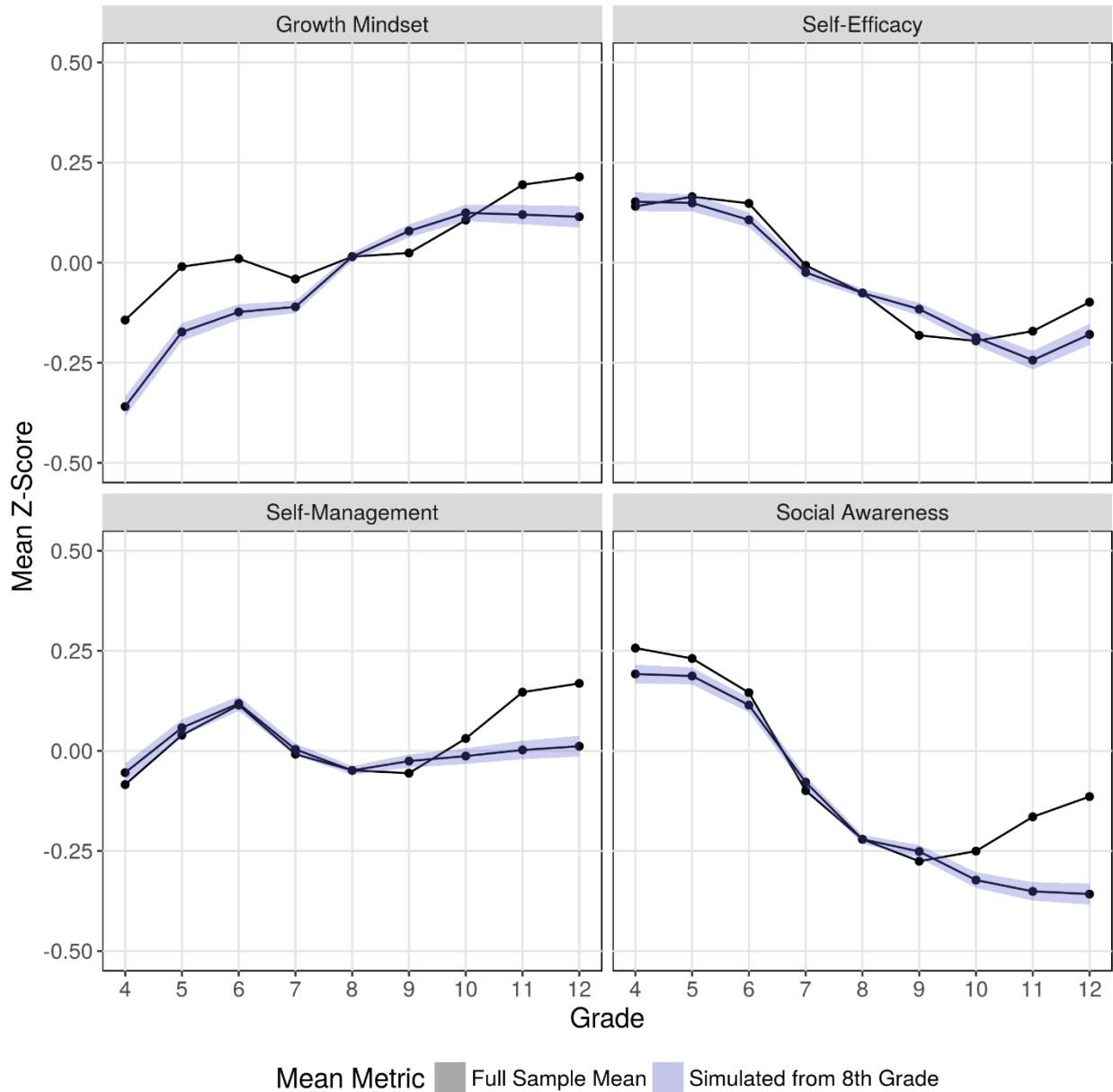
scores, students' SEL does not increase monotonically across grades. Rather, average scores for three of the SEL constructs (growth mindset is the partial exception) decline markedly from Grade 6 to Grade 9 before appearing to recover partially or entirely in high school. In some cases, the differences in average scores across grade levels are substantial in magnitude. For example, the level of social awareness reported by students in Grade 4 is more than 0.5 standard deviations greater than that reported by students in Grade 9. A parallel figure based on cross-sectional data from the 2014-15 school year (not shown) shows nearly identical patterns, suggesting that the patterns evident in Figure 7 reflect consistent differences across grade levels rather than idiosyncratic differences across the cohorts of students enrolled in a grade in a given year.

Figure 7. Mean Social-Emotional Construct Score and Test Score by Grade, 2015-16



We next use the method described above to simulate the developmental trajectory of each SEL construct between Grades 4 and 12. Figure 8 plots these simulated trends for each construct alongside the same grade-level means reported in Figure 7. Because the simulations are anchored at the full-sample mean for Grade 8, the two lines overlap at that point by construction. In addition to simulated means, the figure also displays 95 percent confidence intervals around each simulated data point. These confidence intervals become larger as one moves away from Grade 8 in either direction, due to the compounding of measurement error from combining estimates of gain scores across multiple grade levels. Although the full-sample means are also measured with uncertainty, their confidence intervals are trivially small due to the large number of surveyed students and are excluded to enhance the figure’s legibility.

Figure 8. Mean Social-Emotional Construct Score: Full Sample vs. Simulated Trend Based on 2015-16 Changes



The simulated trends presented in Figure 8 confirm what the cross-sectional means suggested: the social-emotional skills measured by the CORE districts do not increase monotonically as students advance through school. Growth mindset is a partial exception, with students who remain enrolled in CORE districts from Grade 4 to Grade 12 registering fairly steady growth of approximately 0.06 standard deviations annually between Grade 4 and Grade 10, before leveling off through the remainder of high school. Conversely, students' scores on self-efficacy and social awareness decline by 0.38 and 0.56 standard deviations, respectively, between Grade 4 and Grade 12, with the most rapid declines occurring while students are enrolled in the middle-school grades. In the case of self-management, student scores increase

by roughly 0.2 standard deviations between Grade 4 and Grade 6, decline by a similar amount by Grade 8, and remain roughly stable thereafter.

The differences between the simulated and cross-section trends evident in Figure 8 reflect the influence of non-random entry into or exit from the districts participating in the SEL survey across grade levels. For example, the simulation shows students making substantial gains in growth mindset between Grade 4 and Grade 9 that were not clearly evident in the cross-sectional analysis. Substantively, this implies that students reporting higher levels of growth mindset are more likely to exit participating districts prior to Grade 9 (or that students entering in those years tend to report lower levels). It could be the case that students with more of a growth mindset—those who believe that their intelligence can be improved with effort—are more likely to seek out alternatives to traditional public schools during the late elementary and middle school years. As a result, the increases in growth mindset made by students remaining enrolled in CORE district schools evident in the simulated trend do not translate into higher levels of growth mindset across grade levels when examined in a cross-section.

Even more striking are the gaps between the simulated trends and cross-sectional means at the end of high school, as only the latter shows evidence of improvements in each construct after Grade 10. In this case, the cross-sectional differences across grade levels appear to be a consequence of students who score lower on each SEL construct being more likely to leave CORE district high schools prior to Grade 12.

Subgroup Results

Figures 9, 10, and 11 present trends separately by student gender, economic disadvantage, and race/ethnicity, respectively. The SEL scores of female and male students attending CORE district schools differ notably in terms of both levels and trends. As shown in Figure 9, girls exhibit a sizable advantage over boys with respect to both self-management and social awareness that is present across all grade levels but becomes smaller as students age. In the case of self-management, girls score 0.4 standard deviations higher than boys in Grade 4 and 0.2 standard deviations higher in Grade 12; the bulk of the narrowing of the gender gap for self-management occurs between Grade 6 and Grade 8, when girls experience a larger decline in this construct. In the case of social awareness, girls' advantage starts at 0.32 standard deviations in Grade 4 and narrows to 0.15 standard deviations by Grade 12.

Trends for girls and boys differ even more dramatically for self-efficacy. For this construct, girls start out with a modest 0.12 standard deviation over boys in Grade 4. However, girls experience a decline in self-efficacy that is particularly steep between Grade 6 and Grade 8, when girls' self-efficacy scores decline by 0.11 standard deviations by Grade 6 and by an additional 0.46 standard deviations by Grade 11, before recovering modestly in Grade 12. Boys also register a decline in self-efficacy between Grades 6 and 11, but the slope in their scores is far more gradual. As a result of these trends, girls' self-efficacy starts to lag that of boys in Grade 6 and is roughly one-third of a standard deviation lower throughout high school.

In contrast with the other three constructs, trends with respect to growth mindset are quite similar across genders. Girls exhibit a small advantage over boys in elementary school that narrows (and becomes statistically insignificant) in middle school, but reemerges slightly in high school.

Figure 9. Simulated Trends in Mean Social-Emotional Construct Score by Gender, 2015-16

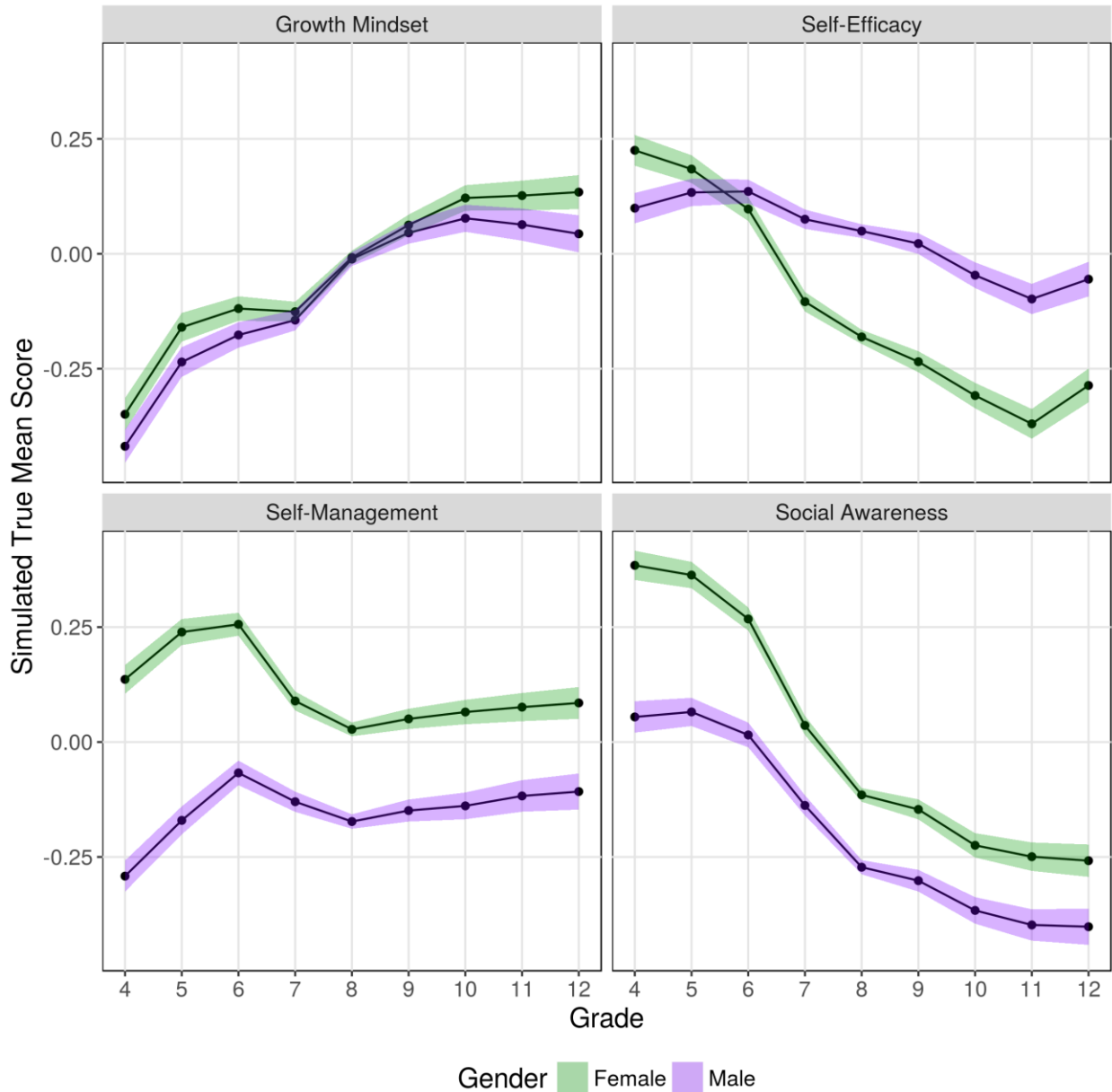
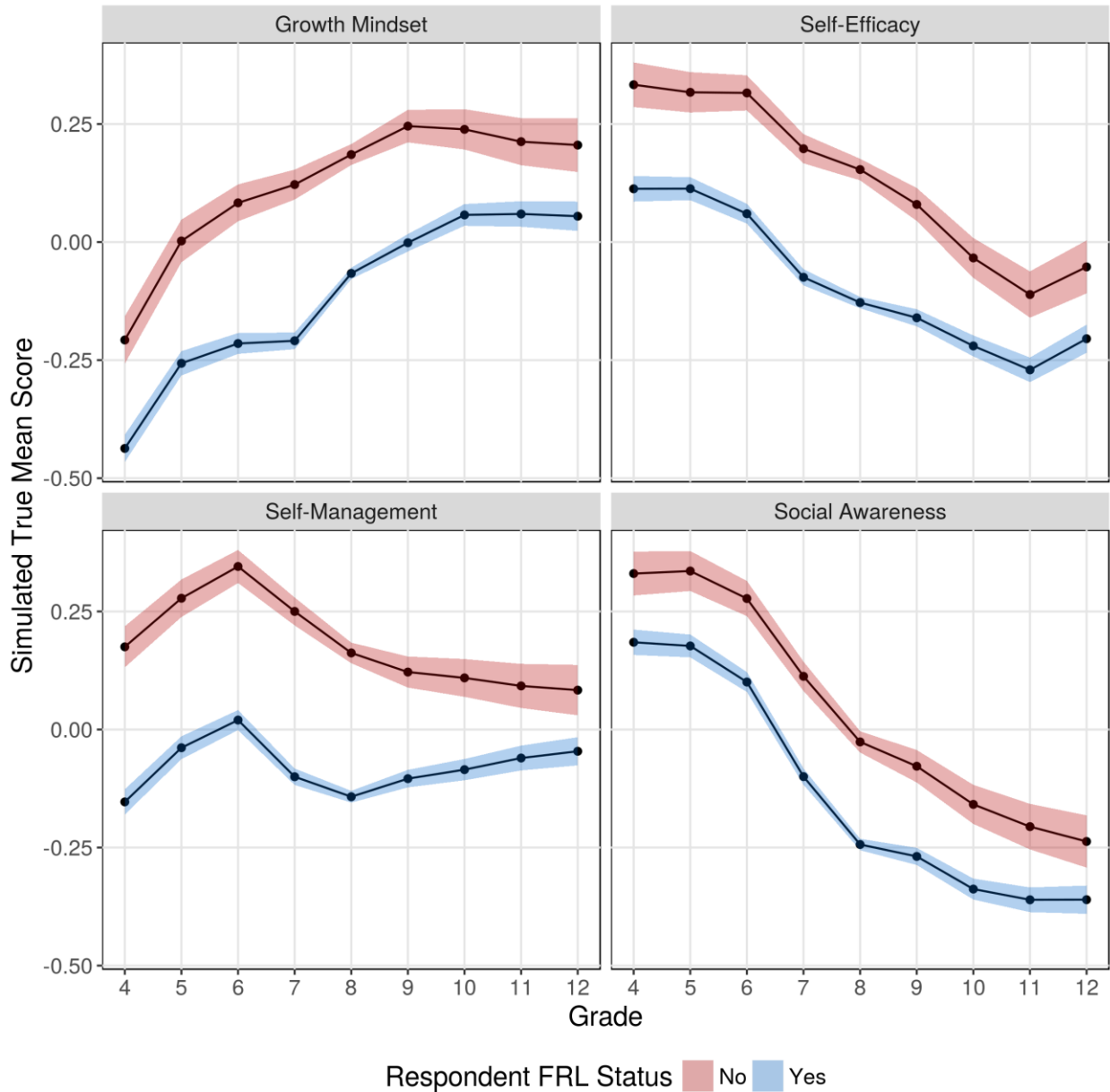


Figure 10 reveals that students who are not economically disadvantaged exhibit higher levels of each SEL construct across all grade levels than those who are. These gaps at Grade 4 range in magnitude from 0.2 standard deviations for social awareness to 0.35 standard deviations for self-management. The gaps widen somewhat in the middle-school grades before narrowing again in high school. In the case of self-management, the result is a narrowing of the

self-management gap associated with economic disadvantage from 0.35 standard deviations in Grade 4 to 0.14 standard deviations in Grade 12. Gaps in growth mindset and self-efficacy also narrow over this eight-year period, but by a smaller amount.

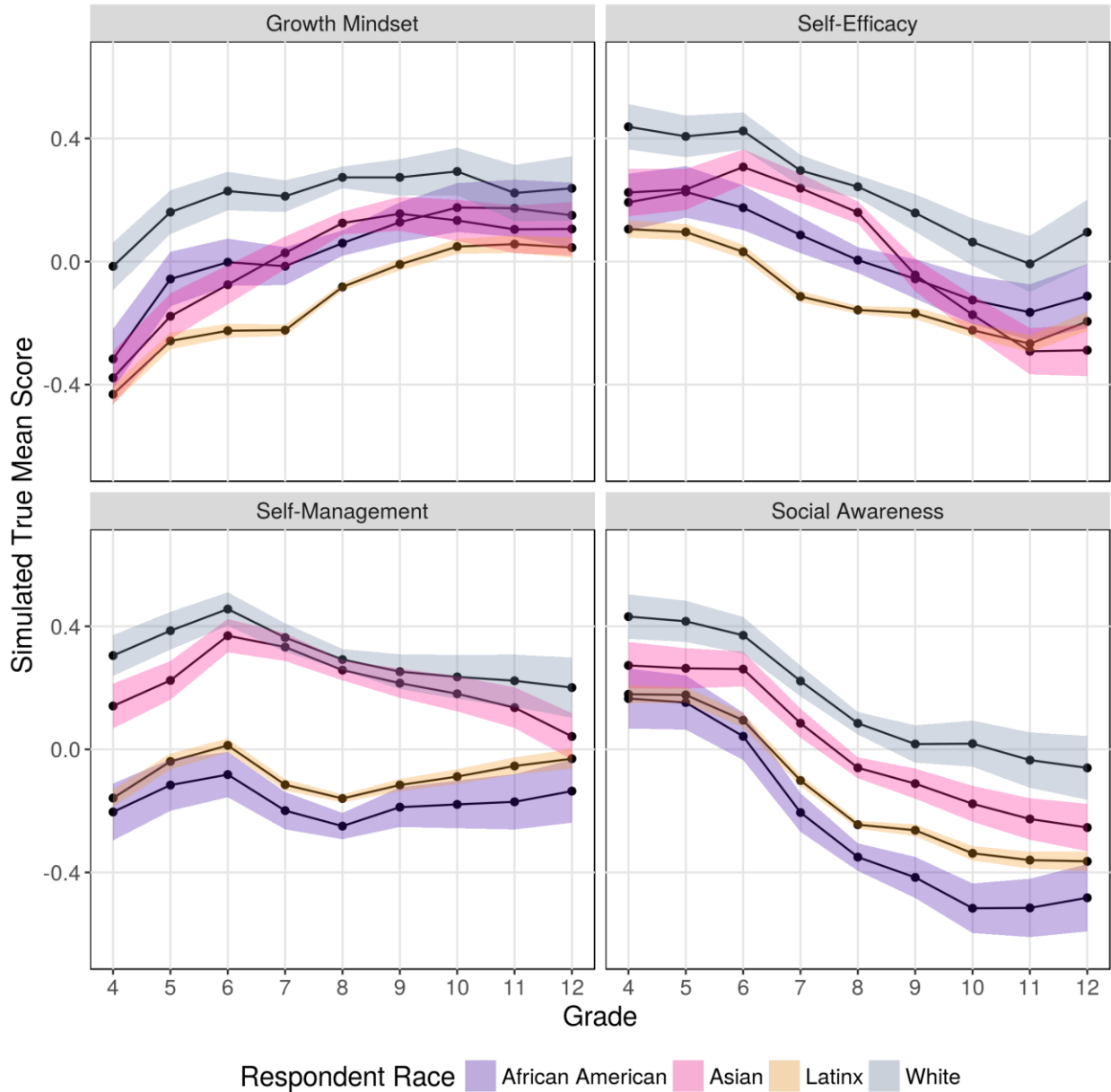
Figure 10. Simulated Trends in Mean Social-Emotional Construct Score by Economic Disadvantage, 2015-16



Finally, Figure 11 documents differences in both levels and trends in SEL for students of different racial and ethnic backgrounds. Consistently across grade levels, White students report higher levels of each SEL construct than do other student groups, though the levels of self-management reported by Asian students follow closely behind those of White students and often do not differ by a statistically significant amount. African American and Latinx students—

the latter comprising the bulk of students enrolled in the CORE districts—generally report lower levels of self-management and social awareness than do White and Asian students. In the case of self-management, these gaps narrow from 0.48 and 0.51 standard deviations in Grade 4 for African American and Latinx students, respectively, to 0.25 and 0.35 standard deviations by Grade 12. However, in the case of social awareness, the size of the gap between African American and Latinx students and White students widens modestly over the same period.

Figure 11. Simulated Trends in Mean Social-Emotional Construct Score by Race/Ethnicity, 2015-16



The patterns observed across racial groups for growth mindset and self-efficacy are more complex. In Grade 4, White students’ growth mindset scores exceed those of each of the

other groups by roughly 0.4 standard deviations. Although growth mindset increases for students of all races by Grade 12, these initial gaps favoring White students narrow by more than half. The timing of the largest increases in growth mindset is delayed for Latinx students, however, with the bulk of the gains occurring between Grade 7 and Grade 9, rather than in elementary school. As a result, Latinx students report noticeably lower levels of growth mindset than all other groups throughout the upper elementary and middle-school grades. In the case of self-efficacy, White, African American, and Latinx students follow similar trends from Grade 4 to Grade 12, with the scores of White students exceeding those of African American and Latinx students by roughly 0.2 and 0.3 standard deviations, respectively. Trends for Asian students are quite different, however. Their self-reported self-efficacy increases through Grade 6 and remains very close to that of White students through Grade 8. However, Asian students' self-efficacy drops by more than one-third of a standard deviation between Grade 8 and Grade 11, leading them to emerge as the lowest-scoring group on this construct by the end of high school.

Discussion

This paper uses self-reported measures from nearly 400,000 students in Grades 4-12 across six school districts to simulate trends over time in four specific social-emotional competencies: Self-management, self-efficacy, growth mindset, and social awareness. In the following, we discuss our findings with respect to our main research questions: (i) How do four key SEL constructs (growth mindset, self-efficacy, self-management, and social awareness) develop from Grade 4-12? And (ii) how do these patterns vary by gender, socioeconomic status, and race/ethnicity?

Trends in SEL

The simulated cohort trends first reveal that, at least within our sample districts, cross-sectional differences in mean SEL scores across grade levels are skewed by non-random entry and exit of students and therefore do not provide an accurate picture of average growth in SEL skills over time. For example, the simulated trends shows students making larger gains with respect to growth mindset in elementary and middle school than would be implied by a naive interpretation of the cross-sectional data. Conversely, the cross-sectional data misleadingly suggest that students make strong gains with respect to each construct in high school, whereas the simulated cohort trends reveal that this pattern is driven at least in part by differential dropout behavior. That is, students with lower levels of these social-emotional competencies are more likely to leave high schools in participating districts between Grades 9 and 12. These differences between the cross-sectional data and simulated trends suggest that the CORE districts' survey-based measures of SEL are valid predictors of continued enrollment in participating districts. They also highlight the importance of the method we employ to examine SEL development.

Our findings substantiate claims that the development of students' does not proceed linearly or even monotonically over time (e.g., Pintrich & Zusho, 2002; Schunk & Meece, 2006; Schunk & Pajares, 2002). For self-management, we find that this construct improves in the late-elementary grades but then decreases during middle school before tapering off in high school.

This aligns with findings from other researchers (e.g., Duckworth et al., 2010; Edossa et al., 2018; Kopp, 1982; Rothbart et al., 2006; Rueda, Posner, & Rothbart, 2005; West et al., 2016), but unlike these studies, we examine self-management trends across both childhood and adolescence in a single dataset. Self-efficacy also appears to decline in middle school in line with previous studies (e.g., Anderman et al., 1999; Pajares & Valiante, 1999; Pintrich & Zusho, 2002; Schunk & Meece, 2006; Schunk & Pajares, 2002; Urdan & Midgley, 2003; Wigfield et al., 2006). As an exception, students' growth mindset seems to increase with age during early adolescence, consistent with the findings of West and colleagues (2016) for three schools, with these gains tapering off while students are in high school.

Notably, however, the results of our simulated cohort trends differ from most existing research on social skills (e.g., Choudhury, et al., 2006; Eccles, 1999; Piaget, 1972; Rubin, et al., 2005; Ryan, 2001; Wigfield et al., 2006; Yurgelun-Todd, 2007). We find that social awareness in fact declines across grades, with a particularly notable plunge between Grades 6 and 9. This discrepancy is potentially due to the particular focus of the CORE districts' measure of social awareness. The CORE items⁴ ask students to quantify how often they engaged in behaviors taking others' thoughts and feelings into account, articulating their own feelings, and getting along with others who are different from them, while other studies measure if peer groups become more central to students' identities and experiences as they age (Rubin, et al., 2005; Ryan, 2001; Wigfield et al., 2006). In addition, third-party observations of students' perspective taking or social skills might increase as students age, even if their self-perceptions of these behaviors or skills may not.

Trends in SEL by Subgroup

Gender. Girls tend to display higher self-management and social awareness than boys across all grade levels, though this gap narrows over time. Although (to our knowledge) this particular trend has not yet been captured in the literature across elementary, middle, and high school, this generally aligns with previous research finding that girls outperform boys in these competencies (e.g., Ablard & Lipshultz, 1998; Duckworth & Seligman, 2006; Gaspar et al., 2018; Kågesten et al., 2016; Moffitt et al., 2011; Pintrich & Zusho, 2002; Wentzel, 1994; Zimmerman & Martinez-Pons, 1990).

Unlike self-management and social awareness, we found little difference in terms of growth mindset between female and male students, which corroborates some existing research (e.g., Macnamara & Rupani, 2017; Storek & Furnham, 2013; Tucker-Drob et al., 2016), even though some scholars have suggested that girls may be more likely to endorse a fixed mindset compared to boys (Dweck, 1986, 2000; Dweck & Simmons, 2014; Halvorson, 2011). Our results indicate that boys and girls report similar levels of growth mindset across elementary, middle, and high school—even as other SEL constructs develop differently.

Indeed, trends in girls' and boys' self-efficacy are particularly striking; although girls report slightly higher self-efficacy in elementary school, their self-efficacy declines far more rapidly than that of boys in middle and high school, before recovering slightly in Grade 12.

⁴ http://coredistricts.org/wp-content/uploads/2017/07/SEL_6.9.17.docx

Although prior studies have detected similar trends in middle school, they also suggest boys and girls display similar self-efficacy in elementary school (Anderman et al., 1999; Wigfield et al., 1991, 1996). Our results indicate that girls exhibit higher self-efficacy in elementary school, implying an even greater decline in self-efficacy for girls over time. In addition, conflicting findings in the field about gender differences may stem from the use of measures of global self-efficacy versus domain-specific self-efficacy (e.g., Pajares, 1996; Pajares & Graham, 1999; Roeser et al., 1996; Smith et al., 2002); our findings suggest that girls and boys display different trends even on measures of global academic self-efficacy.

Socioeconomic status. For students from economically disadvantaged backgrounds, our findings corroborate the broad consensus in the field that students growing up in poverty display lower levels of SEL than their more advantaged peers, and that this trend persists over time (e.g., Bradley & Corwyn, 2002; Brooks-Gunn & Duncan, 1997; Kupersmidt et al., 1995; Takeuchi et al., 1991; Yoshikawa et al., 2012). However, notably, we find that SES-based gaps in SEL tend to narrow in high school, especially for self-management. Importantly, this narrowing does not reflect that students who are economically disadvantaged are more likely to drop out of high school prior to Grade 12, since our analyses include only students who were present in both years of survey data.

Since family and community influences play a pivotal role in SEL (Eccles, 1999), these results suggest that schools, in particular, may have an opportunity to additionally support the development of social-emotional skills among students from economically disadvantaged backgrounds. Given that social-emotional skills are predictive of student's academic achievement and other life outcomes (e.g., Almlund et al., 2011; Heckman et al., 2014), targeted interventions may help to alleviate the detrimental effects of poverty on students' long-term well-being and success.

Race/ethnicity. Both ESSA and California's own Local Control Funding Formula (LCFF) require the disaggregation of student outcomes by race/ethnicity, highlighting the importance of understanding how SEL measures differ among students in these groups. In making such comparisons, however, it is important to keep in mind various factors that could lead students of color to respond to survey items differently than their White peers. Students of color are more likely to be economically disadvantaged and to experience trauma outside of school (Bolger et al., 1995; Chau, Thampi, & Wight, 2010; DeCarlo Santiago, Wadsworth, & Stump, 2011; Hackman, Betancourt, Brodsky, Hurt, & Farah, 2012); these differences not only are risk factors for social and emotional health, but also may make it more difficult for students of color to demonstrate the kinds of behaviors asked about in the survey. Evidence from the CORE districts' school culture and climate survey also reveals that students of color rate their school's culture and climate less favorably than do their White peers, even when they attend the same school (Hough, Kalogrides, and Loeb, 2017); this is consistent with extensive research showing that students' experiences within school differ by race/ethnicity, including well-documented disparities in disciplinary practices and expectations for success (Bankston & Zhou, 2002; Gregory, Skiba, & Noguera, 2010; Lareau & Horvat, 1999; Lewis, 2003; Okonofua, Walton, & Eberhardt, 2016; Tenenbaum & Ruck, 2007; Warikoo & Carter, 2009; Watamura, Phillips, Morissey, McCartney, & Bub, 2011). If these factors are considered when evaluating SEL data,

disaggregating the survey results by race can be useful in prompting educators and other stakeholders in schools serving diverse students to discuss how best to support students of color.

Trends in students' SEL over time look different for students of different racial and ethnic backgrounds, and depending on the construct of interest. We see that White students report higher levels of SEL consistently over time (although Asian students' self-management closely mirrors that of White students). However, the size and trend of the gaps, as well as the group of students affected, vary depending on the construct.

In general, across grades, African American students tend to report lower levels of both self-management and social awareness relative to other racial/ethnic subgroups. Similarly to the findings for socioeconomic disadvantage, gaps between racial/ethnic subgroups appear to widen in middle school but narrow in high school for self-management; however, the gap between White students and both Latinx and African American students increases over time for social awareness.

In general, across grades, Latinx students report lower levels of both growth mindset and self-efficacy compared to other racial/ethnic subgroups overall. For growth mindset, the gap between White students and Asian, African American, and Latinx students narrows substantially from Grade 4 to Grade 12, with Latinx students particularly increasing in growth mindset between Grades 7 and 9.

For self-efficacy, gaps among White, African American, and Latinx students do not change dramatically over time; for Asian students, however, this is not the case. Their self-efficacy initially increases in elementary school, but starts to decline in middle school and continues to plummet in high school, such that they report the lowest levels of self-efficacy by the end of high school of any racial/ethnic subgroup. This aligns with what Eaton and Dembo (1997) found, but provides some indication of how this self-efficacy gap may change across elementary, middle, and high school.

Given the lack of empirical research explicitly examining differences in how students from various racial/ethnic student subgroups develop particular social-emotional competencies (Pintrich & Zusho, 2002), these results provide much-needed preliminary insights into these questions. However, there is still a lot more work to be done. For instance, it may be that these self-report surveys are written, structured, or administered in a way that consistently elicit different kinds of responses from students from different racial/ethnic backgrounds. In addition, the self-report nature of the measures may elicit different kinds of response patterns from different groups of students that might either mask or exacerbate gaps among student subgroups; for instance, White students may be more likely to inflate self-reported measures of SEL relative to students of other races/ethnicities. Ongoing research to examine these possibilities is therefore still needed.

Conclusion

This paper has used the first large-scale panel survey of SEL outcomes to examine trends in students' social and emotional learning across grades, and how those trends differ across student subgroups. We provide new evidence of the validity and reliability of student self-reports of social-emotional skills as measured by the CORE districts' SEL survey, thereby contributing to a growing body of evidence about the appropriate interpretations and uses of these kinds of measures (e.g., Gehlbach & Hough, 2018; Loeb et al., 2018; Meyer et al., 2018). We also apply an innovative method for simulating cohort trends to estimate how students' self-management, growth mindset, self-efficacy, and social awareness develop between grade 4 and grade 12 among students expected to enroll continuously in the six districts participating in the survey in the 2014-15 and 2015-16 school years.

We find that self-efficacy, social awareness, and, to a lesser degree, self-management decrease after Grade 6. These findings corroborate prior evidence that, unlike academic achievement, the development of students' social-emotional skills does not proceed linearly or even monotonically over time (e.g., Schunk & Meece, 2005; Schunk & Pajares, 2002). However, growth mindset does appear to increase as students progress through school. These trends are critical for educators and policymakers to understand as they seek to make sense of patterns observed in their students. Declines across grades might not be alarming, but rather a sign of typical development. It may also be the case that the changes over time reflect changes in normative standards, rather than students' underlying skills. For instance, self-efficacy may decline at least in part because younger students tend to overestimate their capabilities and become more realistic as they mature (Pintrich & Zusho, 2002). Moreover, the contrasting patterns observed for growth mindset compared to the other three constructs illustrate how different constructs within the SEL domain vary in their development over time, indicating that they should be measured and assessed individually.

We also find that the development of social-emotional skills as measured by student self-reports varies across subgroups. Girls initially report higher levels of SEL than boys, but these gender gaps narrow over time, and girls experience a sharp decline in self-efficacy in middle school that leads them to lag behind boys on this construct through the end of high school; economically disadvantaged students consistently report lower SEL; and White students generally report higher SEL than Asian, African American, and Latinx students. These results suggest that schools may have an opportunity to additionally support the development of social-emotional skills among student groups who these data suggest are lagging behind. Nonetheless, it is crucial to keep in mind that there are many reasons these gaps might not reflect true discrepancies in social-emotional competencies. Given the limited evidence on how specific social-emotional skills develop over time for students from different backgrounds (Broda, et al., 2018; Eccles, et al., 1998; Good, Aronson, & Inzlicht, 2003; Walton & Cohen, 2011), more work is needed to understand where particular subgroups are excelling and where they may need additional support at various points in their educational trajectories. The narrowing of racial and ethnic gaps with respect to self-management and growth mindset within the CORE districts is encouraging, but more research is needed to shed light on the interpretation of and factors contributing to those trends.

There are some important caveats to the findings presented here. First, although we find evidence of these measures' reliability and validity, student self-report measures nonetheless have important limitations (Duckworth & Yeager, 2015). Self-report measures are susceptible to a range of possible biases, such as reference bias (West et al., 2016) social desirability bias (Paulhus, 1991). In addition, given the likely multi-dimensionality of the latent constructs underlying social emotional learning, self-report measures may capture only a specific aspect of a given social emotional competency. Second, we can follow the development of individual students' SEL over the course of only a single school year; although we simulate cohort trends from Grade 4 to Grade 12, true longitudinal analyses would provide more precise estimates of development over time. Third, our specific findings are may not generalize to settings outside the California school districts in which the data were gathered.

Nevertheless, the data from the CORE districts SEL survey provide a unique opportunity to assess students' social-emotional development across elementary, middle, and high school, and the results suggest fruitful areas for additional research. As additional school systems gather data on SEL at scale, using either survey-based measures or alternative forms of assessment, it should become possible to learn whether the patterns we document are unique to the CORE districts or hold more generally. It is only by better understanding how SEL typically proceeds over the course of students' educational careers that practitioners, researchers, and policymakers can begin to make sense of the patterns of SEL they see in their own data, for their own students.

References

- Ablard, K. E., & Lipschultz, R. E. (1998). Self-regulated learning in high achieving students: Relations to advanced reasoning, achievement goals, and gender, *Journal of Educational Psychology, 90*, 94-101.
- Alexander, K., & Entwisle, D. (1988). Achievement in the first two years of school: Patterns and processes. *Monograph for Research in Child Development, 53*(2):1-157.
- Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), *Handbook of the economics of education* (Vol. 4) (pp. 1-181). Amsterdam: Elsevier, North-Holland.
- Anderman, E. M., Maehr, M. L., & Midgley, C. (1999). Declining motivation after the transition to middle school: Schools can make a difference. *Journal of Research and Development in Education, 32*, 131-147.
- Bandura, A. (1997). *Self-efficacy: The Exercise of Control*. New York: W. H. Freeman.
- Bankston, C. L., & Zhou, M. (2002). Being well vs. doing well: Self-esteem and school performance among immigrant and nonimmigrant racial and ethnic groups. *International Migration Review, 36*(2), 389–415.
- Berg, J., Osher, D., Same, M. R., Nolan, E., Benson, D., & Jacobs, N. (2017). *Identifying, defining, and measuring social and emotional competencies: Final report*. Washington, DC: American Institutes of Research.
- Blackwell, L. A., 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.
- Blum, R. W., & Libbey, H. P. (2004). School connectedness— Strengthening health and education outcomes for teenagers. *Journal of School Health, 74*, 229–299.
- Bolger, K. E., Patterson, C. J., Thompson, W. W., & Kupersmidt, J. B. (1995). Psychosocial adjustment among children experiencing persistent and intermittent family economic hardship. *Child Development, 66*, 1107–1129.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53*(1), 371-399.
- Britner, S. L., & Pajares, F. (2001). Self-efficacy beliefs, motivation, race, and gender in middle school science. *Journal of Women and Minorities in Science and Engineering, 7*, 271-285.
- Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., & Diemer, M. (2018). Reducing inequality in academic success for incoming college students: A randomized trial of growth mindset and belonging interventions. *Journal of Research on Educational Effectiveness*.
- Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *Children and Poverty, 7*(2), 55-71).
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology, 3*, 296–322.
- Chau, M., Thampi, K., & Wight, V. R. (2010). *Basic facts about low-income children, 2009, children under age 18*. New York: National Center for Children in Poverty. Retrieved from <www.nccp.org/publications/pub_975.html>.

- Choudhury, S., Blakemore, S.-J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience, 1*(3), 165-174.
- Claro, S., Paunesku, D., & Dweck, C. S. (2016). Growth mindset tempers the effects of poverty on academic achievement. *Proceedings of the National Academy of Sciences, 113*(31), 8664-8668.
- Colder, C. R., Mott, J., Levy, S., & Flay, B. (2000). The relation of perceived neighborhood danger to childhood aggression: A test of mediating mechanisms. *American Journal of Community Psychology, 28*(1) 83-103.
- Collaborative for Academic, Social, and Emotional Learning [CASEL]. (2005). *Safe and sound: An educational leader's guide to evidence-based social and emotional learning programs— Illinois edition*. Chicago, IL: Author.
- Conger, R. D., Conger, K. J., & Elder, G. H. (1997). Family economic hardship and adolescent adjustment: Mediating and moderating processes. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of growing up poor* (pp. 288-310). New York: Russell Sage Foundation.
- Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources, 43*(4), 738-782.
- Davis, A. J. (2012). Examining gender and socio-economic status on the emotional intelligence of early adolescents. *PCOM Psychology Dissertations*. Paper 211.
- DeCarlo Santiago, C., Wadsworth, M. E., & Stump, J. (2012). Socioeconomic status, neighborhood disadvantage, and poverty-related stress: Prospective effects on psychological syndromes among diverse low-income families. *Journal of Economic Psychology, 32*(2), 218-230.
- Dee, T. S., & West, M. R. (2011). The non-cognitive returns to class size. *Education Evaluation and Policy Analysis, 33*(1), 23-46.
- Deming, D. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics, 132*(4), 1593-1640.
- DiPerna, J. C., & Elliott, S. N. (1999). The development and validation of the Academic Competence Evaluation Scales. *Journal of Psychoeducational Assessment, 17*, 207-225.
- Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*(1), 198-208.
- Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using hierarchical linear modeling: An illustration predicting achievement from self-control. *Social Psychological and Personality Science 1*(4), 311-317.
- Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44*(4), 237-251.
- Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. (1994). Economic deprivation and early childhood development. *Child Development, 65*, 296-318.
- Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 47-70). New York: Russell Sage Foundation.

- Durlak, J. A., Dymnicki, A. B., Taylor, R., Weissberg, R. P., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development* 82(1), 405-432.
- Dusenbury, L., Newman, J. Z., Weissberg, R. P., Goren, P., Domitrovich, C. E., & Mart, A. K. (2015). Developing a blueprint for preschool to high school education in social and emotional learning: The case for state learning standards. In J. A. Durlak, C. E. Domitrovich, R. P. Weissberg, & T. P. Gullotta (Eds.), *Handbook of social and emotional learning: Research and practice* (pp. 532-548). New York: Guilford Press.
- Dweck, C. S. (1986). Motivational processes affecting learning. *The American Psychologist*, 41(10), 1040–1048.
- Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. New York: Psychology Press.
- Dweck, C. (2006). *Mindset: The new psychology of success*. New York, NY: Random House Publishing Group.
- Dweck, C., & Simmons, R. (2014). Why do women fail? Retrieved from <<http://www.cnn.com/2014/07/29/opinion/dweck-simmons-girls-confidence-failure/>>
- Eaton, M. I., & Dembo, M. H. (1997). Differences in the motivational beliefs of Asian American and non-Asian students, *Journal of Educational Psychology*, 89, 433-440.
- Eccles, J. S. (1987). Gender roles and women's achievement related decisions. *Psychology of Women Quarterly*, 11, 135-172.
- Eccles, J. S. (1999). The development of children ages 6 to 14. *The Future of Children*, 9(2), 30-44.
- Eccles, J. S., Adler, T., & Meece, J. L. (1984). Sex differences in achievement: A test of alternate theories. *Journal of Personality and Social Psychology*, 46, 26-43.
- Eccles, J. S., Lord, S., & Midgley, C. (1991). What are we doing to adolescents? The impact of educational contexts on early adolescents. *American Journal of Education*, 99, 521 – 542.
- Eccles, J. S., Wigfield, A., Midgley, C., & Reuman, D. (1993). Negative effects of traditional middle schools on students' motivation. *Elementary School Journal*, 93, 553 – 574.
- Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 1017-1095). New York: Wiley.
- Edossa, A. K., Schroeders, U., Weinert, S., & Artelt, C. (2018) The development of emotional and behavioral self-regulation and their effects on academic achievement in childhood. *International Journal of Behavioral Development*, 42(2), 192-202.
- Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., Johnson, D. W., & Beechum, N. O. (2012). *Teaching adolescents to become learners: The role of noncognitive factors in shaping school performance: A critical literature review*. Chicago: University of Chicago Consortium on Chicago School Research.
- Gaspar, T., Cerqueira, A., Branquinho, C., & Gaspar de Matos, M. (2018). Dimensions of social and personal skills in children and adolescents: Age and gender differences. *International Journal of Development Research*, 8(1), 18394-18400.

- Gehlbach, H., & 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, Stanford, CA.
- Gestsdottir, S., & Lerner, R. M. (2008). Positive development in adolescence: The development and role of intentional self-regulation. *Human Development, 51*, 202-224.
- Good, C., Aronson, J., & Inzlicht, M. (2003). Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology, 24*, 6145–662.
- Graham, S. (1994). Motivation in African Americans. *Review of Educational Research, 64*, 55-117.
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap two sides of the same coin? *Educational Researcher, 39*(1), 59–68.
- Halvorson, H. G. (2011). The trouble with bright girls. Retrieved from <www.psychologytoday.com/blog/the-science-success/201101/the-trouble-bright-girls>.
- Hatzchristou, C., & Hopf, D. (1996). A multiperspective comparison of peer sociometric status groups in childhood and adolescence. *Child Development, 67*, 1085-1102.
- Hackman, D. A., Betancourt, L. M., Brodsky, N. L., Hurt, H., & Farah, M. J. (2012). Neighborhood disadvantage and adolescent stress reactivity. *Frontiers in Human Neuroscience, 6*, 1-11.
- Heckman, J. J., & Kautz, T. (2013). Fostering and measuring skills: Interventions that improve character and cognition. In J. J. Heckman, J. E. Humphries, & T. Kautz (Eds.). *The myth of achievement tests: The GED and the role of character in American life* (pp. 341-430). Chicago: University of Chicago Press.
- Heckman, J. J., Humphries, J. E., & Kautz, T. (Eds.). (2014). *The myth of achievement tests: The GED and the role of character in American life*. Chicago: University of Chicago Press.
- Hill, J. P., & Lynch, M. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Petersen (Eds.), *Girls at puberty* (pp. 201-228). New York: Plenum.
- Hough, H., Kalogrides, D., & Loeb, S. (2017). *Using surveys of students' social-emotional learning and school climate for accountability and continuous improvement*. Stanford, CA: Policy Analysis for California Education.
- Jacobs, J. E., Lanza, S., Osgood, W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509-527.
- Kågesten, A., Gibbs, S., Blum, R. W., Moreau, C., Chandra-Mouli, V., Herbert, A, et al. (2016) Understanding factors that shape gender attitudes in early adolescence globally: A mixed-methods systematic review. *PLoS ONE, 11*(6): e0157805.
- Kopp, C. B. (1982) Antecedents of self-regulation: A developmental perspective. *Developmental Psychology, 18*(2), 199-214.
- Kupersmidt, J. B., Griesler, P. C., DeRosier, M. E., & Patterson, C. J. (1995). Childhood aggression and peer relations in the context of family and neighborhood factors. *Child Development, 66*, 360-375.
- Lareau, A., & Horvat, E. M. (1999). Moments of social inclusion and exclusion: Race, class, and cultural capital in family–school relationships. *Sociology of Education, 72*(1), 37–53.

- Lewis, A. E. (2003). *Race in the schoolyard: Negotiating the color line in classrooms and communities*. New Brunswick, NJ: Rutgers University Press.
- Loeb, S., Christian, M.S., Hough, H., Meyer, R. H., Rice, A., & West, M.R. (2018). *School effects on social-emotional learning. Findings from the first large-scale panel survey of students*. Policy Analysis for California Education, Stanford, CA.
- Macnamara, B. N., & Rupani, N. S. (2017). The relationship between intelligence and mindset. *Intelligence, 64*, 52-59.
- Malecki, C. K., & Elliott, S. N. (2002). Children's social behaviors as predictors of academic achievement: A longitudinal analysis. *School Psychology Quarterly, 17*(1), 1-23.
- Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to adulthood. *Journal of Educational Psychology, 81*, 417-430.
- Masters, G.N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174.
- McCoy, M. B., Frick, P. J., Loney, B. R., & Ellis, M. L. (1999). The potential mediating role of parenting practices in the development of conduct problems in a clinic-referred sample. *Journal of Child and Family Studies, 8*, 477-494.
- McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53*, 185-204.
- McLeod, J., & Shanahan, M. (1993). Poverty, parenting, and children's mental health. *American Sociological Review, 58*, 351-366.
- Meyer, R. H., Wang, Y. C., & Rice, A. B. (2018). *Measuring students' social-emotional learning among California's CORE districts: An IRT modeling approach*. Policy Analysis for California Education, Stanford, CA.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., & Sears, M. R. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, 108*, 2693-2698.
- Muraki, E. (1992). *A generalized partial credit model: Application of an EM algorithm*. (ETS Research Report 92-06). Princeton, NJ: Educational Testing Service.
- National Research Council (2012). *Education for life and work: Developing transferable knowledge and skills for the 21st Century*. Washington, DC: The National Academies Press.
- Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A vicious cycle: A social-psychological account of extreme racial disparities in school discipline. *Perspectives on Psychological Science, 11*(3), 381-398.
- Ortega, S. T., & Corzine, J. (1990). Socioeconomic status and mental disorders. *Research in Community and Mental Health, 6*, 149-82.
- Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research, 66*, 543-578.
- Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*, 124-139.
- Pajares, F., & Johnson, M. J. (1996). Self-efficacy beliefs in the writing of high school students: A path analysis. *Psychology in the Schools, 33*, 163-175.
- Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*, 426-443.

- Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology, 24*, 390-405.
- Papini, D. R., Farmer, F. F., Clark, S. M., Micka, J. C., & Barnett, J. K. (1990). Early adolescent age and gender differences in patterns of emotional self-disclosure to parents and friends. *Adolescence, 25*(100), 959-976.
- Paulhus, D. L. (1991). Measurement and control of response bias. In Robinson, J. P., Shaver, P. R., Wrightsman, L. S. (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.
- Piaget, J. (1972). *The psychology of intelligence*. Totowa, NJ: Littlefield.
- Pintrich, P. R., & Zusho, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 249-284). San Diego, CA: Academic Press.
- Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology, 88*, 408-422.
- Ross, K. M., & Tolan, P. (2017). Social and emotional learning in adolescence: Testing the CASEL model in a normative sample. *Journal of Early Adolescence*.
- Rothbart, M. K., Posner, M. I., & Kieras, J. (2006). Temperament, Attention, and the Development of Self-Regulation. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 338-357). Malden, NJ: Blackwell.
- Rubin, K. H., Coplan, R., Chen, X., Buskirk, A. A., & Wojslawowicz, J. C. (2005). Peer relationships in childhood. In M. Bornstein & M. Lamb (Eds.), *Developmental science: An advanced textbook* (5th ed., pp. 469-512). Mahawh, NJ: Lawrence Erlbaum.
- Rueda, M. R., Posner, M. I., & Rothbart, M. K. (2005). The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology, 28*(2), 573-594.
- Ryan, A. M. (2001). The peer group as a context for the development of young adolescents' motivation and achievement. *Child Development, 72*, 1135-1150.
- Schunk, D. H., & Meece, J. L. (2006). Self-efficacy development in adolescence. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 71-96). Greenwich, CT: Information Age Publishing.
- Schunk, D. H., & Miller, S. D. (2002). Self-efficacy and adolescents' motivation. In F. Pajares & T. Urdan (Eds.), *Academic motivation of adolescents* (pp. 29-52). Greenwich, CT: Information Age.
- Schunk, D. H. & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation*. San Diego, CA: Academic Press.
- Shell, D. FR., Colvin, C., & Bruning, R. H. (1995). Self-efficacy, attribution, and outcome expectancy mechanisms in reading and writing achievement: Grade-level and achievement-level differences. *Journal of Educational Psychology, 87*(3), 386-398.
- Sinclair, J. J., Pettit, G. S., Harrist, A. W., Dodge, K. A., & Bates, J. E. (1994). Encounters with aggressive peers in early childhood: Frequency, age differences, and correlates of risk

- for behaviour problems. *International Journal of Behavioural Development*, 17, 675–696.
- Smith, L., Sinclair, K. E., & Chapman, E. S. (2002). Students' goals, self-efficacy, self-handicapping, and negative affective responses: An Australian senior school student study. *Contemporary Educational Psychology*, 27, 471-485.
- Sonja, P., Milena, V. Z., Jana, K., & Cirila, P. (2009). Students' social behaviour in relation to their academic achievement in primary and secondary school: Teacher's perspective. *Psihologijske Teme*, 18(1), 55-74.
- Spearman, C. C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271–295.
- Snipes, J., & Tran, L. (2017). *Growth mindset, performance avoidance, and academic behaviors in Clark County School District* (REL 2017-226). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Education Laboratory West. Retrieved from Retrieved from <http://ies.ed.gov/ncee/edlabs>.
- Starfield, B. (1989). Child health care and social factors: Poverty, class, race. *Bulletin of the New York Academy of Medicine*, 65, 299–306.
- Steinberg, L. (2007). Risk taking in adolescence: New perspectives from brain and behavioral science. *Current Directions in Psychological Science* 16(2), 55-59.
- Storek, J., & Furnham, A. (2013). Gender, 'g', and fixed versus growth intelligence mindsets as predictors of self-estimated Domain Masculine Intelligence (DMIQ). *Learning and Individual Differences*, 25, 93-98.
- Takeuchi, D. T., Williams, D. R., & Adair, R. K. (1991). Economic stress in the family and children's emotional and behavioral problems. *Journal of Marriage and Family*, 53, 1031–41.
- Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99(2), 253.
- Tucker-Drob, E. M., Briley, D. A., Engelhardt, L. E., Mann, F. D., & Harden, K. P. (2016). Genetically-mediated associations between measures of childhood character and academic achievement. *Journal of Personality and Social Psychology*, 111(5), 790–815.
- Underwood, M.K. (2004). Gender and peer relations: Are the two gender cultured really all that different? In J.B. Kupersmidt & K.A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention* (pp. 21-36). Washington: American Psychological Association.
- Urda, T., & Midgley, C. (2003). Changes in the perceived classroom goal structure and pattern of adaptive learning during early adolescence. *Contemporary Educational Psychology*, 28, 524-551.
- Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science*, 331(6023), 1447–1451.
- Warikoo, N., & Carter, P. (2009). Cultural explanations for racial and ethnic stratification in academic achievement: A call for a new and improved theory. *Review of Educational Research*, 79(1), 366–394.

- Watamura, S. E., Phillips, D. A., Morrissey, T. W., McCartney, K., & Bub, K. (2011). Double jeopardy: Poorer social-emotional outcomes for children in the NICHD SECCYD experiencing home and childcare environments that confer risk. *Child Development, 82*(1), 48–65.
- Wentzel, K.R. (1994). Relations of social goal pursuit to social acceptance, and perceived social support. *Journal of Education Psychology, 86*, 173-182.
- 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., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2016). Promise and paradox: Measuring non-cognitive traits of students and the impact of schooling. *Educational Evaluation and Policy Analysis, 38*(1), 148-170.
- Wigfield, A., Byrnes, J. P., & Eccles, J. S. (2006). Development during early and middle adolescence. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd Ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wigfield, A., Eccles, J., Maclver, D., Reuman, D., & Midgley, C. (1991). Transitions at early adolescence: Changes in children’s domain-specific self-perceptions and general self esteem across the transition to junior high school. *Developmental Psychology, 27*, 552-565.
- Wigfield, A., Eccles, J. S., & Pintrich, P. R. (1996). Development between the ages of 11 and 25. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 148-185). New York: Macmillan.
- Wills, T. A., McNamara, G., & Vaccaro, D. (1995). Parental education related to adolescent stress-coping and substance use: Development of a mediational model. *Health Psychology, 14*, 464–478.
- Winer, A. C., & Thompson, R. A. (2013). *How poverty and depression impact a child's social and emotional competence*. Center for Poverty Research, University of California-Davis. Available at <<https://poverty.ucdavis.edu/policy-brief/how-poverty-and-depression-impact-childs-social-and-emotional-competence>>
- Yeager, D., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., et al. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*(3), 374-391.
- Yoshikawa, H., Aber, J. L., & Beardslee, W. R. (2012). The effects of poverty on the mental, emotional, and behavioral health of children and youth. *American Psychologist, 67*(4), 272-284.
- Yurgelun-Todd, D. (2007) Emotional and cognitive changes during adolescence. *Current Opinion in Neurobiology, 17*, 251-257.
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*, 51-59.

Appendix

Table A1: Descriptive Statistics for Matched Sample

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Grade 12	Total
Male	50.7%	50.3%	50.1%	50.3%	50.0%	50.7%	50.4%	49.3%	49.1%	50.1%
Asian	4.1%	5.0%	7.7%	8.3%	7.8%	7.6%	8.2%	8.9%	9.2%	7.3%
African American	7.1%	7.5%	7.1%	6.7%	6.9%	6.3%	7.1%	7.1%	6.9%	7.0%
Latinx	75.3%	73.4%	71.1%	69.9%	70.6%	73.7%	72.4%	71.0%	70.6%	72.1%
ELL	50.7%	49.9%	43.1%	34.4%	31.3%	26.5%	23.3%	18.6%	17.8%	34.4%
SWD	10.8%	11.9%	11.5%	11.1%	10.5%	10.8%	9.6%	9.2%	10.0%	10.7%
Econ Dis	78.5%	77.4%	77.9%	76.8%	76.4%	78.5%	76.1%	75.0%	76.6%	77.1%
Foster	0.8%	0.8%	0.6%	0.6%	0.5%	0.5%	0.5%	0.4%	0.4%	0.6%
Homeless	1.3%	1.9%	2.1%	2.4%	2.6%	2.1%	2.4%	1.9%	1.8%	2.1%
Total Students	32,975	39,367	35,145	33,568	34,255	29,133	29,285	25,314	23,825	282,867
Total Schools	529	714	379	210	265	259	237	229	225	1,033